

Non-linear and Graphical Methods for Fish Stock Analysis with Statistical Modeling



Tatsuro Akamine*

*National Research Institute of Fisheries Science
Fisheries Research Agency
Fukuura 2-12-4, Yokohama 236-8648, Japan*

Abstract

Useful methods for growth curve fitting, body-size composition analysis, and estimation of population size in fish stocks are presented. These methods are statistically based on the maximum likelihood method and the likelihood ratio test. Mathematical explanation of the standard Richards growth formula with seasonal change, the generalized reproduction model, and the Awaya method for estimating implicit function models are given. Mathematical proofs of the iteration method, called the Hasselblad method, or the EM algorithm for estimating the mixture of normal distributions, and the Marquardt method for general optimization are shown. For population size estimation, the Petersen method for mark–recapture experiments, the quadrat method, and the DeLury removal method are discussed. These are based on the binomial distribution and the classical Bayesian statistical methods which are also discussed. Mathematical proofs of the sum formulae of the binomial and hyper-geometric distributions are given. The virtual population analysis using mortality rates, the Leslie matrix model, and the linear programming for discrete fishing models are also explained. All the methods stated here can be easily carried out using spread-sheet software.

Received on March 4, 2008
Revised on March 13, 2009
Accepted on April 28, 2009
Online published on July 28, 2009

Keywords

- body-size composition
- growth curve
- population size
- reproduction
- survival

1. Introduction

There are many methods to estimate the parameters of fish stocks. Before the 1980s, most of them were approximations, linear regression methods, and generally had low precision. Since the 1980s, the author has developed various non-linear methods by using microcomputers. Nowadays, these methods can be carried out using spread-sheet software. In particular, its graphical functions are useful to understand the statistical models and check parameter values.

In Section 2, the standard growth formula of fish is presented. This formula is based on the Richards growth formula and expanded for seasonal growth change. The non-linear regression method to estimate and test parameters, the generalized reproduction model, and the Awaya method for estimating implicit function models are also explained. In Section 3, the Hasselblad method which estimates parameters of a mixture of normal distributions is explained by using the EM algorithm and the K–L information quantity. The proof of convergence, the Marquardt method and the Jacobi method are also shown. In Section 4, the estimation methods for population size are detailed. These are the Petersen method for mark–recapture experiments, the quadrat

*Corresponding author at:
National Research Institute of Fisheries Science
Fisheries Research Agency
Fukuura 2-12-4, Yokohama 236-8648, Japan
e-mail: akabe@affrc.go.jp

method, and the DeLury removal method. Their statistical models are based on the binomial or hyper-geometric distribution. The classical Bayesian statistical methods and the sum formulae of the binomial or hyper-geometric distributions are explained. Survival models are shown in Section 5. The virtual population analysis (VPA) using mortality rates, the Leslie matrix model related to VPA, and the linear programming method for discrete fishing models are explained.

The above-mentioned models are simple and easy to understand. Statistical methods of point and interval estimations are the maximum likelihood method and the likelihood ratio test. We can easily calculate and draw figures for statistical estimation by using spread-sheet software.

2. Standard growth formula in fish population dynamics

In fish population dynamics, many growth formulae and their expansions are used. Among them, von Bertalanffy growth formula is the most commonly used. For parameter estimation, the linear regression method (e.g. Ford/Walford plot) for data measured with a constant interval, has previously been widely used. However, it was impossible to estimate three parameters at the same time and the curve fitting was not good. In the 1980s, personal computers and BASIC made it possible to estimate non-linear models and obtain good curve-fitting. In this section, some expanded models, the standard growth formula, and estimation method will be shown.

2-1. Traditional growth formulae

In fish population dynamics, the von Bertalanffy growth formulae are widely used for body length (VBGF1)

$$l(t) = l_{\infty} (1 - e^{-k(t-t_0)}) \quad (2.1)$$

and for body weight (VBGF2)

$$w(t) = w_{\infty} (1 - e^{-k(t-t_0)})^3, \quad (2.2)$$

where l and w are the body-sizes, t is the time, k is the growth coefficient, l_{∞} and w_{∞} are the body-sizes at $t = \infty$, t_0 is the extrapolated time when $l = 0$ or $w = 0$. These are solutions of the differential equation

$$\frac{dw}{dt} = \eta w^{2/3} - \kappa w, \quad w = al^3, \quad (2.3)$$

where η , κ and a are constants. On the other hand, the logistic growth formula (LGF)

$$w(t) = \frac{w_{\infty}}{1 + e^{-k(t-c)}}, \quad (2.4)$$

and the Gompertz growth formula (GGF)

$$w(t) = w_{\infty} \exp(-e^{-k(t-c)}), \quad (2.5)$$

where c is the time of the inflection point, are also used in some cases.

Richards (1959) solved the differential equation

$$\frac{dw}{dt} = \eta w^m - \kappa w \quad (2.6)$$

and obtained the growth formula (RGF)

$$w(t) = \left[\frac{\eta}{\kappa} - \left(\frac{\eta}{\kappa} - w_0^{1-m} \right) e^{-(1-m)\kappa t} \right]^{1/(1-m)} \quad (2.7)$$

or

$$w(t) = \left[w_{\infty}^{1-m} - (w_{\infty}^{1-m} - w_0^{1-m}) e^{-(1-m)\kappa t} \right]^{1/(1-m)} \quad (2.8)$$

with the initial condition $(t, w) = (0, w_0)$. This formula corresponds to LGF when $m = 2$, VBGF1 when $m = 0$, VBGF2 when $m = 2/3$, and converges to GGF when $m \rightarrow 1$, $m \neq 1$. However, this convergence is not easy to understand because Eq. (2.6) is an exponential function when $m = 1$ (Taylor 1962).

In ecology and agriculture, France and Thornley (1984) defined RGF as follows:

$$w(t) = \frac{w_0 w_f}{[w_0^n + (w_f^n - w_0^n) e^{-kt}]^{1/n}}, \quad (2.9)$$

where $w_f = w_\infty$. This is the solution of the differential equation

$$\frac{dw}{dt} = \frac{kw(w_f^n - w^n)}{nw_f^n}. \quad (2.10)$$

However, parameters are different from the traditional growth formulae used in fish population dynamics.

For use in fish population dynamics, Pauly and David (1981) suggested the following formula:

$$L_t = L_\infty (1 - e^{-kD(t-t_0)})^{1/D}, \quad D > 0. \quad (2.11)$$

This is the generalized VBGF. Although Schnute (1981) defined RGF as

$$Y(t) = y_\infty \left(1 + \frac{1}{p} e^{-g(t-t_0)}\right)^{-p}, \quad p > 0, \quad (2.12)$$

this is the generalized LGF which is incorporated in RGF. He also defined the generalized VBGF

$$Y(t) = y_\infty (1 - e^{-g(t-t_0)})^p, \quad p > 0 \quad (2.13)$$

and GGF

$$Y(t) = y_\infty \exp(-e^{-g(t-t_0)}). \quad (2.14)$$

These four equations are included in RGF. The application and potential of RGF has not been sufficiently understood in fish population dynamics.

2-2. Standard formula of RGF in fish population dynamics

Akamine (1988b) defined the differential equation of RGF as follows:

$$\frac{dw}{dt} = kw \frac{1 - (w/w_\infty)^r}{r}, \quad (2.15)$$

where r is the parameter which decides the shape of the growth curve. It is easy to understand the convergence when $r \rightarrow 0$ as follows:

$$\frac{dw}{dt} \rightarrow kw \ln \left(\frac{w_\infty}{w} \right) = kw (\ln w_\infty - \ln w) \quad (2.16)$$

by using the definition of the logarithm

$$\ln y = \lim_{r \rightarrow 0} \frac{y^r - 1}{r}. \quad (2.17)$$

Equation (2.16) is the differential equation of GGF. The transformation

$$u = \frac{(w_\infty/w)^r - 1}{r} \quad (2.18)$$

reduces Eq. (2.15) to the following differential equation

$$\frac{du}{dt} = -\frac{w_{\infty}^r}{w^{r+1}} \frac{dw}{dt} = -ku. \quad (2.19)$$

The solution of this with the initial condition $(t, w) = (c, w_{\infty}/(1+r)^{1/r})$ is

$$w(t) = \frac{w_{\infty}}{(1 + re^{-k(t-c)})^{1/r}}. \quad (2.20)$$

This is the standard formula of RGF used in fish population dynamics, which also includes generalized VBGF, GGF and generalized LGF (Fig. 1). The initial condition means the inflection point of this curve. We can understand that Eq. (2.20) converges to Eq. (2.5) when $r \rightarrow 0$ by using the following definition of the exponential:

$$e^x = \lim_{r \rightarrow 0} (1 + rx)^{1/r}. \quad (2.21)$$

When $r = -p < 0$, this standard formula can be rewritten as

$$w(t) = w_{\infty} (1 - pe^{-k(t-c)})^{1/p} = w_{\infty} (1 - e^{-k(t-t_0)})^{1/p}, \quad (2.22)$$

$$t_0 = c + \frac{\ln p}{k}.$$

This is the generalized VBGF. By using the allometric function

$$w' = aw^b, \quad (2.23)$$

RGF is transformed to

$$w'(t) = \frac{aw_{\infty}^b}{(1 + re^{-k(t-c)})^{b/r}} = \frac{w'_{\infty}}{(1 + r'e^{-k(t-c')})^{1/r'}}, \quad (2.24)$$

$$w' = aw_{\infty}^b, \quad r' = \frac{r}{b}, \quad c' = c + \frac{\ln b}{k}.$$

This is also RGF.

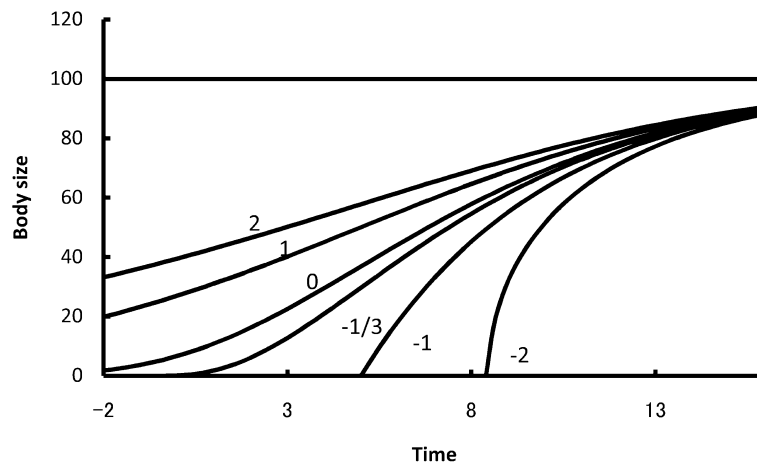


Fig. 1. Graphs of the Richards growth formulae. Figures are values of r . Revised from *Bull. Japan Sea Natl. Fish. Res. Inst.*, **38**, Fig. 2, Akamine T. Estimation of parameters for Richards model. 187–200, 1988, with permission from Japan Sea National Fisheries Research Institute, Fisheries Research Agency.

Schnute (1981) defined his new growth formula as the differential equations

$$\frac{dY}{dt} = YZ, \quad \frac{dZ}{dt} = -Z(a + bZ). \quad (2.25)$$

Equation (2.15) can be rewritten as

$$\frac{dw}{dt} = wZ, \quad Z = k \frac{1 - (w/w_\infty)^r}{r}. \quad (2.26)$$

These are modified as follows:

$$\frac{dZ}{dt} = -k \frac{w^{r-1}}{w_\infty^r} \frac{dw}{dt} = -k \left(\frac{w}{w_\infty} \right)^r Z = -Z(k - rZ). \quad (2.27)$$

Therefore, Eq. (2.25) is equal to Eq. (2.15) which is RGF itself.

2-3. Seasonal growth formula

Pitcher and MacDonald (1973) presented the growth formulae (PM1)

$$L_{t_g} = L_\infty (1 - e^{-k(t_g - t_0)}), \quad \text{when } \cos\left(\frac{2\pi t_s}{52}\right) < \text{sw} \quad \text{then } \frac{dt_g}{dt_s} = 0, \quad (2.28)$$

and (PM2)

$$L_{t_r} = L_\infty (1 - e^{-k_1}), \quad k_1 = C \sin\left[\frac{2\pi(t_r - s_1)}{52}\right] + k(t_r - t_0), \quad (2.29)$$

where t_r is real time (weeks), t_g is growth time, $t_s = t_r - s$, s is the starting point for cosine, sw is the threshold value for cosine, C is the magnitude constant, and s_1 is the starting point for sine. PM1 is the switching model and PM2 is the sine curve model. And they define the water temperature θ as

$$\theta = m + q \cos\left[\frac{2\pi(t_r - s_2)}{52}\right], \quad (2.30)$$

where m is the mean water temperature, q is the magnitude constant, and s_2 is the time lag between temperature and growth switch. Using this function, they rewrote the growth formulae as follows (PM1):

$$\text{When } \theta = \text{sw}_1, \quad \text{then } \frac{dt_g}{dt_s} = 0, \quad (2.31)$$

and (PM2):

$$k_1 = Cq \sin\left[\frac{2\pi(t_r - s_3)}{52}\right] + k(t_r - t_0), \quad (2.32)$$

where sw_1 is the threshold temperature, and s_3 is the time lag between mean temperature and growth oscillation. Pauly and Gashutz (1979) presented the equation (PG)

$$L_t = L_\infty \left\{ 1 - \exp\left[-k(t - t_0) + \frac{Ck}{2\pi} \sin 2\pi(t - t_s)\right] \right\}. \quad (2.33)$$

This is simpler than PM2. Haddon (2001) used the formula (HD)

$$L_t = L_\infty (1 - e^{-k}), \quad (2.34)$$

$$k = C_1 \sin\left[\frac{2\pi(t - s_1)}{52}\right] + C_2 \sin\left[\frac{2\pi(t - t_s)}{p}\right] + k(t - t_0).$$

This is a more complex model than PM2 and PG. Appeldoorn (1987) suggested that these formulae have a problem as $L(t_0) \neq 0$.

Cloern and Nichols (1978) defined the differential equation

$$\frac{dL}{dt} = L'_{\max} - b(t)(L - L_{\max}), \quad b(t) = a_1 \left\{ 1 + \sin \left[\frac{\pi}{180}(t + \theta) \right] \right\}, \quad (2.35)$$

where L'_{\max} is maximal growth rate when body size is minimal, and L_{\max} is the upper limit of body size. They obtained the solution (CN)

$$L(t) = L_{\max} - (L_{\max} - L_{\min}) \exp \left\{ -a_1(t - t_0) - \frac{180a_1}{\pi} \left[\cos \frac{\pi(t_0 + \theta)}{180} - \cos \frac{\pi(t + \theta)}{180} \right] \right\}. \quad (2.36)$$

Although this model overcomes the problem stated by Appeldoorn, this solution has two further problems. The first is that the parameter L_{\min} is not necessary and the second is that the amplitude of the periodic function is fixed to be 1.

Hoening and Hanumara presented the formula in 1982 (HH)

$$L_t = L_{\infty} - L_{\infty}e^x, \quad (2.37)$$

$$x = -k(t - T_0) - \frac{kC}{2\pi} \sin 2\pi(t - T_2) + \frac{kC}{2\pi} \sin 2\pi(T_0 - T_2).$$

Furthermore, Somers (1988) presented the formula (SO)

$$L_t = L_{\infty} (1 - e^{-[k(t-t_0)+S(t)-S(t_0)]}), \quad S(t) = \frac{Ck}{2\pi} \sin 2\pi(t - t_s). \quad (2.38)$$

This is a simpler version of HH. These models also overcome the problem stated by Appeldoorn. In summary, researchers in fish population dynamics have previously used the following expression for exponential:

$$k(t - t_0) + f(t). \quad (2.39)$$

However, the meaning of the trigonometric function $f(t)$ is difficult to understand.

2-4. Standard formula for seasonal growth

Akamine (1986) solved the differential equation

$$\frac{dl}{dt} = k(l_{\infty} - l)f(t), \quad f(t+1) = f(t), \quad (2.40)$$

and obtained the growth formula (AK1)

$$l(t) = l_{\infty} (1 - e^{-k[F(t)-F(t_0)]}), \quad \frac{dF}{dt} = f(t). \quad (2.41)$$

This is the simplest expression for seasonal growth of VBGF1. He used the functions

$$f(t) = \frac{1+a}{2} + \frac{1-a}{2} \cos 2\pi(t - t_1), \quad (2.42)$$

$$F(t) = \frac{1+a}{2}t + \frac{1-a}{4\pi} \sin 2\pi(t - t_1). \quad (2.43)$$

Where $a \leq f(t) \leq 1$. In the case that $f(t)$ is the water temperature, $F(t)$ means the cumulative water temperature. When $a < 0$, this formula includes negative growth. He also obtained an expanded expression of the logistic growth formula (AK2)

$$l(t) = \frac{l_{\infty}}{1 + e^{-k[F(t)-F(c)]}}, \quad (2.44)$$

and that of the Gompertz growth formula (AK3)

$$l(t) = l_{\infty} \exp \left(-e^{-k[F(t)-F(c)]} \right). \quad (2.45)$$

Further, Akamine (1988b) obtained an expanded Richards growth formula (RA):

$$w(t) = \frac{w_{\infty}}{\left(1 + r e^{-k[F(t)-F(c)]}\right)^{1/r}}. \quad (2.46)$$

This is the standard formula for seasonal growth. Akamine (1993, 1995c) showed the general case as follows: Let the differential equation be

$$\frac{dw}{dt} = g(w)f(t). \quad (2.47)$$

He later modified this to

$$\frac{dw}{g(w)} = f(t)dt = dF. \quad (2.48)$$

The integral of this is

$$G(w) = F(t) + C, \quad (2.49)$$

where C is an integral constant. Then he obtained the general solution

$$w = G^{-1}(F(t) + C), \quad (2.50)$$

and the particular solution

$$w = G^{-1}(F(t) - F(c)). \quad (2.51)$$

Therefore, we can get the expanded formula with the operation $t \rightarrow F(t)$.

He used the following standardization:

$$\int_0^1 f(t)dt = F(1) - F(0) = 1 \quad (2.52)$$

and functions

$$f(t) = 1 + A \cos 2\pi(t - t_1), \quad (2.53)$$

$$F(t) = t + \frac{A}{2\pi} \sin 2\pi(t - t_1). \quad (2.54)$$

Equation (2.43) is rewritten as

$$kF(t) = \frac{1+a}{2}k \left[t + \frac{1-a}{1+a} \frac{1}{2\pi} \sin(t - t_1) \right]. \quad (2.55)$$

In this formula, the parameter a strongly influences the growth coefficient k . Therefore, the following operations are needed for comparison of parameter values:

$$k \leftarrow \frac{1+a}{2}k, \quad A = \frac{1-a}{1+a}. \quad (2.56)$$

Kiso *et al.* (1992) applied the following function to data for the Masu salmon (*Oncorhynchus masou*):

$$F(t) = t + \frac{A}{2\pi} \sin 2\pi(t - t_1) + \frac{C}{6\pi} \sin 6\pi(t - t_1). \quad (2.57)$$

The standard formula (2.46) can include many growth formulae. For example, it includes the switching model PM1. Let

$$f(t) = \begin{cases} \frac{1}{1 - (t_3 - t_2)} & (0 \leq t < t_2, \quad t_3 \leq t < 1), \\ 0 & (t_2 \leq t < t_3), \end{cases} \quad (2.58)$$

then we obtain

$$F(t) = \begin{cases} n + \frac{t - n}{1 - (t_3 - t_2)} & (n \leq t < n + t_2), \\ n + \frac{t_2}{1 - (t_3 - t_2)} & (n + t_2 \leq t < n + t_3), \\ n + \frac{t - n - (t_3 - t_2)}{1 - (t_3 - t_2)} & (n + t_3 \leq t < n + 1). \end{cases} \quad (2.59)$$

This is equal to PM1. Furthermore, Pauly *et al.* (1992) proposed a new switching model (PE):

$$L_t = L_\infty (1 - e^{-q}), \quad (2.60)$$

$$q = k(t' - t_0) + \frac{k}{Q} [\sin Q(t' - t_s) - \sin Q(t_0 - t_s)], \quad Q = \frac{2\pi}{1 - \text{NGT}},$$

where NGT is the duration of non-growth within a year. This is also included in the standard formula (2.46). Let

$$f(t) = \begin{cases} 1 + \cos 2\pi \frac{t - t_1}{1 - T} & (0 \leq t < t_2), \\ 0 & (t_2 \leq t < t_3), \\ 1 + \cos 2\pi \frac{t - T - t_1}{1 - T} & (t_3 \leq t < 1), \end{cases} \quad (2.61)$$

$$t_2 = t_1 + \frac{1 - T}{2}, \quad t_3 = t_1 + \frac{1 + T}{2}, \quad \int_0^1 f(t) dt = 1 - T, \quad T = \text{NGT},$$

then we obtain

$$F(t) = \begin{cases} n(1 - T) + S(t - n) & (n \leq t < n + t_2), \\ n(1 - T) + S(t_2) & (n + t_2 \leq t < n + t_3), \\ n(1 - T) + S(t - n - T) & (n + t_3 \leq t < n + 1), \end{cases} \quad (2.62)$$

$$S(t) = t + \frac{1 - T}{2\pi} \sin 2\pi \frac{t - t_1}{1 - T}.$$

This is equal to PE.

2-5. Parameter estimation and statistical test of growth formulae

For parameter estimation of growth formulae, Pitcher and MacDonald (1973) calculated the fitting by hand and the direct-search method by computer. They also used the multivariable regression to obtain values of C , k and t_0 in Eq. (2.29) as follows:

$$\ln \left(1 - \frac{L_{t_r}}{L_\infty} \right) = -C \sin \left[\frac{2\pi(t_r - s_1)}{52} \right] - kt_r + kt_0. \quad (2.63)$$

Pauly and Gashutz (1979) also used the multiple regression to obtain values of C , k , t_0 and t_s in Eq. (2.33) as follows:

$$\begin{aligned}\ln\left(1 - \frac{L_t}{L_\infty}\right) &= -k(t - t_0) + \frac{Ck}{2\pi} \sin 2\pi(t - t_s) \\ &= kt_0 - kt + \left(\frac{Ck}{2\pi} \cos 2\pi t_s\right) \sin 2\pi t \\ &\quad - \left(\frac{Ck}{2\pi} \sin 2\pi t_s\right) \cos 2\pi t.\end{aligned}\quad (2.64)$$

These are linear regression models that cannot obtain a high precision estimate for L_∞ .

Schnute (1981) used the simplex search method, which was transformed to microcomputers, for his growth model. The objective functions are the least-squares method with no weighting for raw data, and for their logarithm values.

Akamine (1982) used the Gauss–Seidel method in BASIC to estimate parameters of a mixture of normal distributions and Akamine (1984) used the Marquardt method for the same model. These objective functions are the least-squares method with no weighting. Akamine (1986) used the Marquardt method for the estimation of the growth curves and the objective function is the weighted least-squares method.

When we measure fish body-size w at time t_i , let n_i be the number of specimens, \bar{w}_i be the mean, and σ_i^2 be the variance. When $t = t_i$ the residual sum of squares is

$$\begin{aligned}\epsilon^2 &= \sum_{j=1}^n \frac{\{w_j - w(t)\}^2}{\sigma^2} \\ &= \sum \frac{\{w_j - \bar{w} + \bar{w} - w(t)\}^2}{\sigma^2} \\ &= \sum \frac{(w_j - \bar{w})^2}{\sigma^2} + 2\{\bar{w} - w(t)\} \sum \frac{w_j - \bar{w}}{\sigma^2} + n \frac{\{\bar{w} - w(t)\}^2}{\sigma^2} \\ &= (n - 1) + 0 + \frac{\{\bar{w} - w(t)\}^2}{\sigma^2/n},\end{aligned}\quad (2.65)$$

where the subscript i is omitted, w_j is the body-size of the j -th individual at time t_i , and $w(t)$ is the growth formula. Therefore, we can calculate the residual sum of squares only by using the mean \bar{w}_i and its variance. According to the central limit theorem, the distribution of \bar{w}_i is the normal distribution and its variance is given

$$\delta_i^2 = \frac{\sigma_i^2}{n_i}.\quad (2.66)$$

Thus, the objective function of the weighted least-squares method is defined as

$$Y = \sum_{i=1}^k \frac{\{\bar{w}_i - w(t_i)\}^2}{\delta_i^2}\quad (2.67)$$

The optimization method which minimizes this objective function gives the parameter values of the growth formula.

(Example 1) Fitting the growth formula to clam data.

The clam data of the Oita prefecture is shown in Table 1. Although these data have the means and their standard deviations, they have no number of specimens. Therefore, let $\delta \equiv \sigma$ and use the weighted least-squares method (2.67). By using the optimization method of the spread-sheet software (Gorie 2001), we obtain the Gompertz formula

$$w(t) = 69.80 \exp\left(-e^{-0.7348(t-2.038)}\right).$$

Table 1. The growth data of the clam (*Meretrix lusoria*). SL is the Shell length (mm). Data were excerpted from Kamijoh *et al.* Hamaguri no shigen baiyou gijutu kaihatu kenkyu houkokusho, 1985, with permission from Fisheries Research Institute, Oita Prefectural Agriculture, Forestry and Fisheries Research Center.

Age	SL	S.D.	Age	SL	S.D.
0.615	5.4	1.7	3.615	53.1	1.9
0.835	6.5	1.4	3.835	54.8	2.6
1.615	17.6	3.8	4.200	55.2	3.8
1.835	21.3	1.9	4.369	56.2	3.1
2.200	21.3	3.3	4.615	60.0	2.7
2.369	28.1	5.1	4.835	61.9	3.1
2.615	41.5	3.9	5.200	62.2	3.2
2.835	43.4	3.7	5.369	62.5	2.5
3.200	43.2	3.6	5.615	65.6	2.1
3.369	48.0	3.0			

From this solution as the initial values, we obtain the expanded growth formula for seasonal growth

$$w(t) = 68.57 \exp \left(-e^{-0.7881[F(t)-F(2.277)]} \right),$$

$$F(t) = t + \frac{1.497}{2\pi} \sin 2\pi(t - 0.4187).$$

Because $A > 1$, this growth formula has periods of negative growth. This is not valid for hard structures such as shells. These data are for samples taken only four times in a year. It requires more samples for a better estimation. Figure 2 is given by the graphical function of spread-sheet software. This curve has periods of little negative growth because software interpolates. It seems more natural for the real growth pattern than the exact drawing.

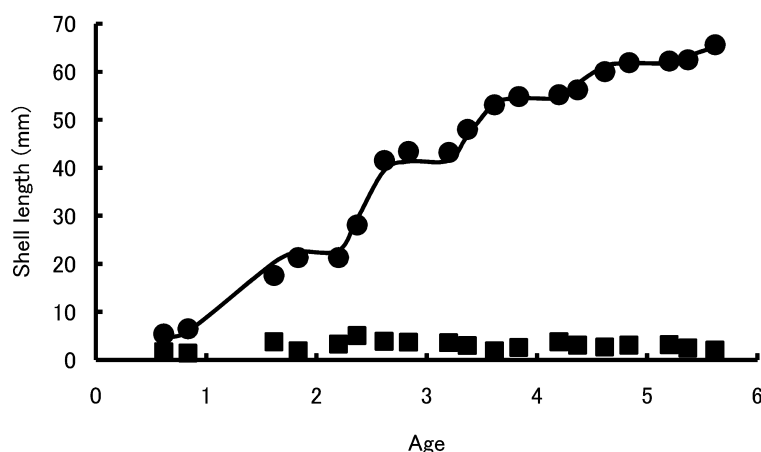


Fig. 2. The growth curve of the clam. Black circles are SL, and squares are S.D. Revised with permission from Akamine T. Taicho-osei data karano nenrei to seicho no suitei. In: *Shigen Hyoka notameno Suuchi Kaiseki*, Suisangaku Series 66, p. 127, Fig. 8.6, Kouseisha Kouseikaku, Tokyo, 1987. ©1987, the Japanese Society of Fisheries Science.

The interval estimation and the test of the growth formula are easy because the distribution of Y is the χ^2 distribution (Akamine 2004). The null hypothesis

$$H_0 : \theta_j = \theta_{j0} \quad (j = 1, \dots, p), \quad (2.68)$$

where p is the number of parameters to estimate, gives the following distribution:

$$Y_0 - Y_{\min} \sim \chi^2(p), \quad (2.69)$$

where Y_0 is the objective function of the true values and Y_{\min} is that of the point estimate. We can obtain the confidence area of parameters by using this distribution. On the other hand, the AIC (Akaike Information Criterion) is defined as

$$AIC = -2 \ln L_{\max} + 2p, \quad (2.70)$$

where $\ln L_{\max}$ is the maximum log likelihood. In this case, the AIC is rewritten as

$$AIC = Y_{\min} + 2p + \text{constant}. \quad (2.71)$$

When there is no weight, the minimum of the residual squares

$$S = \sum_{i=1}^m (w_i - w(t_i))^2, \quad (2.72)$$

where m is the number of specimens, gives the optimum parameter values. As same as for the linear regression, we can use the F distribution for the interval estimation or the test as follows:

$$\frac{(S_0 - S_{\min})/p}{S_{\min}/(m - p)} \sim F(p, m - p), \quad (2.73)$$

where S_0 is the residual squares of the true values and S_{\min} is that of the point estimate. On the other hand, the AIC is approximated in this case as

$$AIC = m \ln S_{\min} + 2p + \text{constant}. \quad (2.74)$$

The example of the model selection by using the AIC is shown in Kiso *et al.* (1992).

As a typical case, the test of the growth difference between male and female is shown as follows: Let the number of male data be m_M , that of female data be m_F . The null hypothesis

$$H_0 : \theta_{jM} = \theta_{jF} \quad (j = 1, \dots, p) \quad (2.75)$$

is used for this test. For the weighted case, we can use the following distribution:

$$Y_{M+F} - Y_M - Y_F \sim \chi^2(p). \quad (2.76)$$

On the other hand, for the non-weighted case, we can use

$$\frac{(S_{M+F} - S_M - S_F)/p}{(S_M + S_F)/(m_M + m_F - 2p)} \sim F(p, m_M + m_F - 2p). \quad (2.77)$$

Although the confidence area of the linear model is an ellipse, that of the non-linear one is a banana-shape. In fisheries science, the wrong discussion has sometimes been seen in which the overlapping between male and female confidence intervals means that the same growth curve can be fitted. Because the confidence interval does not mean the probability in which the true value exists, it is better to use the null hypothesis for these subjects.

(Example 2) Fitting VBGF1 for Pacific hake data.

The data is shown in Table 2. This data has no variance. Thus, F test (2.77) is adequate. The optimization method of the spread-sheet software gives two VBGF1s for male and female by using Eq. (2.72) as follows (Fig. 3):

$$\begin{aligned} l_M(t) &= 55.98(1 - e^{-0.3856(t-0.1713)}), & S_M &= 19.42, \\ l_F(t) &= 61.23(1 - e^{-0.2962(t+0.05726)}), & S_F &= 28.80. \end{aligned}$$

On the other hand, the total data of male and female gives

$$l_{M+F}(t) = 59.29(1 - e^{-0.3205(t-0.01043)}), \quad S_{M+F} = 79.77$$

Thus, Eq. (2.77) is

$$F = \frac{(79.77 - 48.22)/3}{48.22/(24 - 6)} = 3.925.$$

This is significant at 5%, but not at 1%, because $F(3, 18) = 3.160$ (5%), 5.902 (1%).

The confidence region of female growth parameters can be estimated using Eq. (2.73), which is rewritten as

$$S_0 = S_{\min} \left(1 + \frac{p}{m-p} F(p, m-p) \right). \quad (2.78)$$

We can obtain the confidence region of parameters by using S_0 . The null hypothesis must be

$$H_0 : k = k_0, \quad l_{\infty} = l_{\infty 0}, \quad t_0 = t_{00}. \quad (2.79)$$

In this case $p = 3$, $m = 13$, $F(3, 10) = 3.708$ (5%), $S_{\min} = S_F = 28.8$, and we obtain $S_0 = 60.9$ for the condition. By using the optimization method of spread-sheet software, we obtain the minimum $k = 0.200$ ($l_{\infty} = 66.3$, $t_0 = -0.761$), and the maximum $k = 0.413$ ($l_{\infty} = 58.3$, $t_0 = 0.346$), where these values are edges of the three-dimensional confidence region. These VBGF1s are shown as thick lines in Fig. 4. In general, there is a positive correlation between k and t_0 , negative correlations between k and l_{∞} , and l_{∞} and t_0 in VBGF1. The minimum and maximum estimates of l_{∞} and t_0 are omitted.

Table 2. The body-length data (cm) of the Pacific hake (*Merluccius productus*). Data were excerpted with permission from *Fishery Bulletin*, 73, Dark TA. Age and growth of Pacific hake, *Merluccius productus*. 336–355, 1975.

Age	Female	Male
1	15.40	15.40
2	28.03	26.93
3.3	41.18	42.23
4.3	46.20	44.59
5.3	48.23	47.63
6.3	50.26	49.67
7.3	51.82	50.87
8.3	54.27	52.30
9.3	56.98	54.77
10.3	58.93	56.43
11.3	59.00	55.88
12.3	60.91	—
13.3	61.83	—

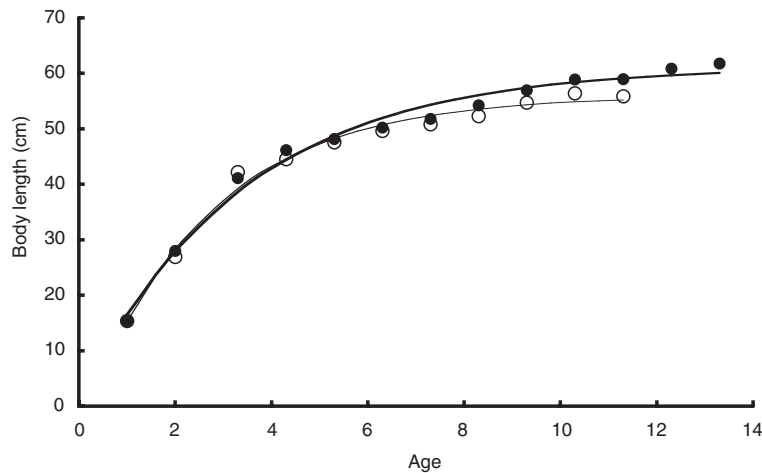


Fig. 3. Growth curves of the Pacific hake. Black circles and the thick line are female data and white circles and the thin line are male. Redrawn with permission after *Fishery Bulletin*, 77, Fig. 1, Kimura DK. Likelihood methods for the von Bertalanffy growth curve. 765–775, 1980.

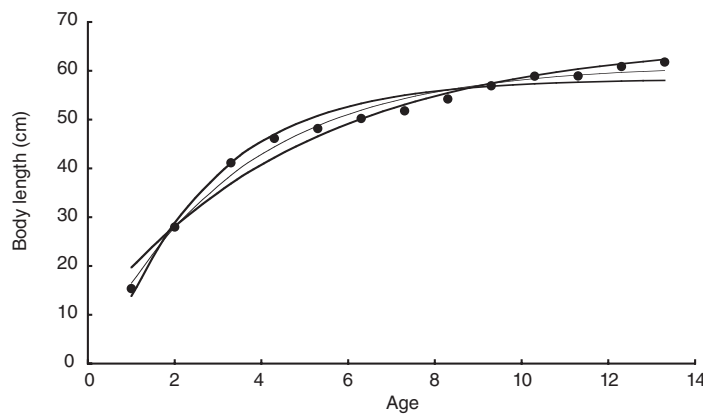


Fig. 4. The 95% confidence interval of the growth curve for female Pacific hake. Two thick lines show growth curves for the minimum and maximum k . Redrawn with permission after Akamine T, *Suisan Shigen Kaiseki no Kiso*, p. 32, Fig. 2.6. Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku Co., Ltd.

2-6. The generalized reproduction model

The treatment of the reproduction model is similar to that of the growth formula. A generalized reproduction model was proposed by Schnute (1985):

$$R = \alpha S(1 - \beta \gamma S)^{1/\gamma}. \quad (2.80)$$

This model can be rewritten as

$$R = \frac{\alpha S}{(1 + r\beta S)^{1/r}}, \quad (2.81)$$

where S is the spawning biomass, R is the recruit biomass, and α , β and r are the parameters which determine the shape of the curve. When $r = 1$, Eq. (2.81) is equal to the Beverton–Holt model

$$R = \frac{\alpha S}{1 + \beta S}. \quad (2.82)$$

When $r \rightarrow 0$, Eq. (2.81) converges to the Ricker model

$$R = \alpha S e^{-\beta S}. \quad (2.83)$$

When $r = -1$, Eq. (2.81) is equal to the parabola which is the Shaeffer model of the surplus production model. When $r \rightarrow \infty$, Eq. (2.81) converges to a straight line (Fig. 5). The gradient at the origin is α . These curves have the following characteristics:

- (1) When $r < 0$, they have an intersection point or a contact point on the abscissa at $S = -1/(r\beta)$. The right side over an intersection point or a contact point of these reproduction models has no meaning.
- (2) When $r < 1$, they have a maximal point at $S = 1/[(1-r)\beta]$.
- (3) When $-1 < r < 1$, they have an inflection point at $S = 2/[(1-r)\beta]$.

Formerly, we obtain this model as follows: The survival equation is

$$\frac{dN}{dt} = - \left[1 + bS \left(\frac{N}{S} \right)^r \right] N, \quad (2.84)$$

where N is the number of fish. The solution of this equation with the initial condition $N(0) = S$ is

$$N(t)^r = \frac{e^{-rt} S^r}{1 + b(1 - e^{-rt})S}. \quad (2.85)$$

Substitution of $R = N(t)$, $\alpha = e^{-t}$ and $\beta = b(1 - e^{-rt})/r$ into this equation gives Eq. (2.81). On the other hand, the following differential equation gives the same solution (Akamine 1994):

$$\frac{dR}{dS} = R \left[\frac{1}{S} - \beta \left(\frac{R}{\alpha S} \right)^r \right]. \quad (2.86)$$

The transformation

$$u = \frac{R}{\alpha S} \quad (2.87)$$

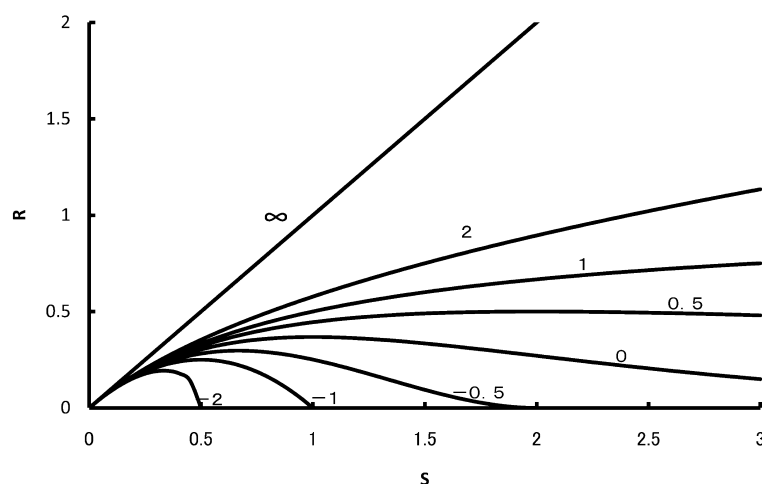


Fig. 5. The graphs of the generalized reproduction model. Figures are values of r . Redrawn with permission after Schnute J. A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sci.*, **42**(3), 419–429, Fig. 2, 1985. ©2008, NRC Canada.

gives the equation

$$\frac{du}{dS} = \frac{1}{\alpha S} \frac{dR}{dS} - \frac{R}{\alpha S^2} = -\beta u^{r+1}. \quad (2.88)$$

The solution of this with the initial condition $R'(0) = \alpha$ is Eq. (2.81).

2-7. Parameter estimation for reproduction model

In the case of estimating R from S , it is better to generally use the regression theory. If we consider both errors of R and S , the weighted least-squares method with errors in both axes is applicable as follows: Let the reproduction model be the implicit model

$$f(S, R, \alpha, \beta) = 0. \quad (2.89)$$

The objective function with m -paired data $(s_i, r_i, \delta_{s_i}^2, \delta_{r_i}^2)$:

$$Y = \sum \left[\frac{(s_i - S_i)^2}{\delta_{s_i}^2} + \frac{(r_i - R_i)^2}{\delta_{r_i}^2} \right], \quad f(S_i, R_i, \alpha, \beta) = 0, \quad (2.90)$$

is adequate for this estimation (Awaya 1991). We can minimize Eq. (2.90) by using Lagrange's undetermined multiplier method.

Awaya (1983) applied the Newton method for the implicit function model. Let the implicit function be

$$f(t, w, \theta) = 0, \quad (2.91)$$

where t is an independent variable, w is a subordinate variable, and θ are parameters. This function cannot be rewritten as $w = g(t, \theta)$. Let the data be m sets of (t_i, μ_i, δ_i^2) and the objective function be

$$Y = \sum_{i=1}^m \frac{(\mu_i - w_i)^2}{\delta_i^2}, \quad f(t_i, w_i, \theta) = 0. \quad (2.92)$$

We can obtain adequate parameter values by minimizing this objective function. Let the objective function be expanded as follows:

$$Y^+ = \sum_i [v_i(\mu_i - w_i)^2 + \lambda_i f(t_i, w_i, \theta)], \quad (2.93)$$

where $v_i = 1/\delta_i^2$, and λ_i is the undetermined multiplier. Thus, the following simultaneous equations can be obtained:

$$\frac{\partial Y^+}{\partial w_i} = -2v_i(\mu_i - w_i) + \lambda_i \frac{\partial f_i}{\partial w_i} = 0 \quad (i = 1, \dots, m), \quad (2.94)$$

$$\frac{\partial Y^+}{\partial \theta_j} = \sum_i \lambda_i \frac{\partial f_i}{\partial \theta_j} = 0 \quad (j = 1, \dots, n), \quad (2.95)$$

$$\frac{\partial Y^+}{\partial \lambda_i} = f_i = 0 \quad (i = 1, \dots, m), \quad (2.96)$$

where $f_i = f(t_i, w_i, \theta)$, and n is the number of parameters. Let

$$\Delta w_i = w_i - w_{i0}, \quad (2.97)$$

$$\Delta \theta_j = \theta_j - \theta_{j0}, \quad (2.98)$$

where the subscript 0 means the initial values. The linear approximation of Eq. (2.96) is

$$f_i = f_{i0} + \frac{\partial f_{i0}}{\partial w_i} \Delta w_i + \sum_j \frac{\partial f_{i0}}{\partial \theta_j} \Delta \theta_j = 0, \quad (2.99)$$

where $f_{i0} = f(t_i, w_{i0}, \theta_0)$. Thus, the following equations can be obtained:

$$\Delta w_i = -\xi_i \left(f_{i0} + \sum_j \frac{\partial f_{i0}}{\partial \theta_j} \Delta \theta_j \right), \quad \xi_i = \left(\frac{\partial f_{i0}}{\partial w_i} \right)^{-1}. \quad (2.100)$$

On the other hand, substituting Eq. (2.97) into Eq. (2.94) with $f_i = f_{i0}$ gives

$$2v_i(d_{i0} - \Delta w_i) - \lambda_i \xi_i^{-1} = 0, \quad (2.101)$$

where $d_{i0} = \mu_i - w_{i0}$. Substituting Eq. (2.100) into Eq. (2.101) gives

$$\lambda_i = 2\xi_i v_i \left[d_{i0} + \xi_i \left(f_{i0} + \sum_j \frac{\partial f_{i0}}{\partial \theta_j} \Delta \theta_j \right) \right]. \quad (2.102)$$

Substituting these into Eq. (2.95) with $f_i = f_{i0}$ gives the following equations:

$$\sum_i \left(\frac{\partial f_{i0}}{\partial \theta_k} \xi_i v_i \xi_i \sum_j \frac{\partial f_{i0}}{\partial \theta_j} \Delta \theta_j \right) = - \sum_j \frac{\partial f_{i0}}{\partial \theta_k} \xi_i v_i \{d_{i0} + \xi_i f_{i0}\} \quad (k = 1, \dots, n). \quad (2.103)$$

These are the normal equations of the least-squares method for the implicit function model.

In practice, if we first decide the initial values of θ_i , the initial values of w_i , which satisfy Eq. (2.96), can be obtained by using numerical calculation. Solving Eq. (2.103) for the modified vector of parameters $\Delta \theta_i$ gives the modified vector of variables Δw_i from Eq. (2.100) and new values of f_i from Eq. (2.99). Therefore, solving Eq. (2.103) again and repeating these procedures gives the solution. The linear approximation errors of f_i converge to 0. This Awaya method is regarded as the Newton method for the multivariable implicit function model. Although its convergence area is narrow, it converges in 10 times or so. This method can be modified to the Marquardt method which has a larger convergence area (Awaya 1991; Akamine 1999). The merit of the Awaya method is that it requires only one time to calculate the correct value of w_i .

2-8. Supplement

In fish population dynamics, allometry

$$w = al^b \quad (2.104)$$

is used for the relationship between the body-weight and body-length. Nowadays, we can estimate body-weight from body-length data by using the following two regression models:

$$w = al^b + \epsilon, \quad (2.105)$$

$$\ln w = \ln a + b \ln l + \epsilon, \quad (2.106)$$

where ϵ is the residual. Each objective function is

$$Y_1 = \sum (w - al^b)^2, \quad (2.107)$$

$$Y_2 = \sum (\ln w - \ln a - b \ln l)^2. \quad (2.108)$$

Y_2 is better than Y_1 because Y_2 is expected to have the same variance for each data. On the other hand, the objective function (2.90) is better to estimate parameter values a and b when both w and l have errors. The objective function must be selected for the purpose.

Ohnishi and Akamine (2006) presented the growth formula for shells (OA). The differential equation of this is

$$\frac{dw}{dt} = \alpha s - \beta w - \gamma \frac{ds}{dt}, \quad (2.109)$$

where $s = \xi l^2$, $w = \eta l^3$. Thus we obtain the equation

$$\frac{dl}{dt} = \frac{kl(l_\infty - l)}{l + a}. \quad (2.110)$$

OA is equal to VBGF1 when $a = 0$ and converges to LGF when $a \rightarrow \infty$. Although OA is similar to RGF, it does not include GGF and VBGF2. The integral of OA is

$$\left(\frac{l(t)}{l_0}\right)^A \left(\frac{l_\infty - l_0}{l_\infty - l(t)}\right)^{A+1} = e^{kt}, \quad (2.111)$$

where $A = a/l_\infty$, $l_0 = l(0)$. This is an implicit model, so the Awaya method is useful for parameter estimation.

3. Analysis of the body-size composition

Estimating the age composition from the body-size composition data is a basic method in fish population analysis. Before the 1960s, graphical methods for a mixture of normal distributions were widely used. Hasselblad (1966) invented the iteration method for this model by using the maximum likelihood method. Akamine (1982, 1984, 1985, 1987b, 1988a) proposed BASIC programs for the Gauss–Seidel method and the Marquardt method by using the maximum likelihood method and the least-squares method. By use of spread-sheet software, the Hasselblad method is easy to apply. In this section, the Hasselblad method is explained by using the EM algorithm. The Jacobi method and the Marquardt method are also explained.

3-1. Statistical model

The body-size composition is regarded as a mixture of normal distributions which is defined as

$$G(x) = \sum_{i=1}^n p_i g_i(x), \quad (3.1)$$

where x is the body-size, subscript i is the number of each distribution, and $p_i > 0$ is a mixture ratio which has the condition

$$\sum_{i=1}^n p_i = 1. \quad (3.2)$$

Function $g_i(x) \geq 0$ is a normal distribution:

$$g_i(x) = N(x, \mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2} \frac{(x - \mu_i)^2}{\sigma_i^2}\right]. \quad (3.3)$$

The estimates of p_i , μ_i and σ_i^2 can be statistically obtained from the body-size composition data. The condition (3.2) decreases the freedom of parameters. Thus, we have to estimate $(3n - 1)$ parameters at the same time.

The maximum likelihood method is adequate to estimate parameters of the probability distribution. When the body-size data is obtained as x_1, \dots, x_m , the likelihood for this model is defined as

$$L = \prod_x G(x) = G(x_1) \cdots G(x_m). \quad (3.4)$$

The parameter values which give the maximum likelihood are the most adequate estimation. In practice, the log likelihood is used for the objective function as follows:

$$Y^* = \ln L = \sum_x \ln G(x) = \ln G(x_1) + \cdots + \ln G(x_m) < 0. \quad (3.5)$$

In the case of the histogram for the data, let $H^*(x)$ be the frequency of a histogram, and $H(x) > 0$ be the relative frequency

$$H^*(x) = mH(x), \quad \sum_x H(x) = 1, \quad (3.6)$$

where x is the class mark of the body size. Thus, the log likelihood is defined as

$$Y^* = \sum_x H^*(x) \ln G(x) = mY, \quad Y = \sum_x H(x) \ln G(x) < 0. \quad (3.7)$$

We can obtain the maximum Y by using the iteration method for computers.

3-2. Hasselblad method

Hasselblad (1966) defined the mixture ratio as follows:

$$p_n = 1 - \sum_{i=1}^{n-1} p_i. \quad (3.8)$$

In this case, the condition of the solution

$$\frac{\partial Y}{\partial p_i} = 0 \quad (3.9)$$

is rewritten as

$$\sum_x h_i(x) = \sum_x h_n(x), \quad (3.10)$$

where

$$h_i(x) = \frac{H(x)g_i(x)}{G(x)}. \quad (3.11)$$

By multiplying p_i to both sides of Eq. (3.10), he obtained

$$p_i \sum_x h_i(x) = p_i \sum_x h_n(x). \quad (3.12)$$

The sum of both sides are

$$\sum_{i=1}^n \left(p_i \sum_x h_i(x) \right) = \sum_x H(x) = 1, \quad (3.13)$$

$$\sum_{i=1}^n \left(p_i \sum_x h_n(x) \right) = \sum_x h_n(x). \quad (3.14)$$

Therefore,

$$\sum_x h_n(x) = 1. \quad (3.15)$$

The resulting equation from Eqs. (3.12) and (3.15) is

$$p_i = p_i \sum_x h_i(x). \quad (3.16)$$

Thus, he obtained the following iteration method:

$$p_i^{(t+1)} = \left(p_i \sum_x h_i(x) \right)^{(t)}. \quad (3.17)$$

The right upper (t) means the t -th iteration value. Substituting old values in the right side provides new values in the left side.

On the other hand, the condition of the solution:

$$\frac{\partial Y}{\partial \mu_i} = 0 \quad (3.18)$$

is rewritten as

$$\sum_x h_i(x)(x - \mu_i) = 0. \quad (3.19)$$

Thus, he obtained the following iteration:

$$\mu_i^{(t+1)} = \left(\frac{\sum_x h_i(x)x}{\sum_x h_i(x)} \right)^{(t)}. \quad (3.20)$$

In the same way, the condition of the solution

$$\frac{\partial Y}{\partial \sigma_i^2} = 0 \quad (3.21)$$

is rewritten as

$$\sum_x h_i(x) [\sigma_i^2 - (x - \mu_i)^2] = 0. \quad (3.22)$$

Thus, he obtained the iteration

$$\sigma_i^{2(t+1)} = \left[\frac{\sum_x h_i(x)(x - \mu_i)^2}{\sum_x h_i(x)} \right]^{(t)}. \quad (3.23)$$

Although the Hasselblad method is stable and useful, it is difficult to fully understand the properties of iteration (3.17).

3-3. Undetermined multiplier method

Akamine and Matsumiya (1992) obtained Eq. (3.17) by using Lagrange's undetermined multiplier method. Let the objective function be

$$Y^+ = Y - \lambda \left(\sum_{i=1}^n p_i - 1 \right), \quad (3.24)$$

where λ is an undetermined multiplier. The condition of the solution

$$\frac{\partial Y^+}{\partial p_i} = 0 \quad (3.25)$$

is rewritten as

$$\sum_x \frac{H(x)g_i(x)}{G(x)} = \lambda. \quad (3.26)$$

The value of λ is determined as

$$\lambda = \lambda \sum_i p_i = \sum_x \sum_i \frac{H(x)p_i g_i(x)}{G(x)} = \sum_x H(x) = 1. \quad (3.27)$$

Thus, we obtain the following non-linear equations:

$$\sum_x h_i(x) = 1. \quad (3.28)$$

By multiplying p_i to both sides of this equation, we obtain Eq. (3.16) and iteration (3.17).

3-4. EM algorithm

The Hasselblad method is regarded as an application of the EM algorithm which has E-step for estimating the missing data and M-step for maximizing the likelihood to estimate parameters. The EM algorithm is developed for the model which has random missing data. Although the model of a mixture of normal distributions has no missing data, it has the latent data as follows:

Define the latent data $z_i(x)$ as the relative frequency of age i and size x , and $q_i(x)$ as the probability of age i in the size x as follows:

$$q_i(x) = \frac{p_i g_i(x)}{\sum_i p_i g_i(x)} = \frac{p_i g_i(x)}{G(x)}. \quad (3.29)$$

In E-step, $z_i(x)$ can be estimated as the expectation:

$$z_i(x) = H(x)q_i(x) = \frac{H(x)p_i g_i(x)}{G(x)}. \quad (3.30)$$

These equations are natural and the calculation amount of E-step is zero.

For the optimization of parameters in M-step, there are defined relations

$$\sum_x \sum_i z_i(x) = \sum_x H(x) = 1, \quad \sum_x \sum_i p_i g_i(x) = \sum_x G(x) = 1. \quad (3.31)$$

We obtain the following log likelihood:

$$Q = \sum_x \sum_i z_i(x) \ln(p_i g_i(x)) = \sum_x \sum_i z_i(x) \ln p_i + \sum_x \sum_i z_i(x) \ln g_i(x). \quad (3.32)$$

To maximize this Q -function for the parameter p_i , let the objective function be

$$Q^+ = Q - \lambda \left(\sum_i p_i - 1 \right). \quad (3.33)$$

The condition of the solution

$$\frac{\partial Q^+}{\partial p_i} = 0 \quad (3.34)$$

is rewritten as

$$\sum_x \frac{z_i(x)}{p_i} = \lambda. \quad (3.35)$$

The value of λ is calculated as

$$\lambda = \lambda \sum_i p_i = \sum_x \sum_i z_i(x) = 1. \quad (3.36)$$

Thus, we obtain the iteration

$$p_i^{(t+1)} = \sum_x z_i(x)^{(t)}. \quad (3.37)$$

This is the same as Eq. (3.17). For the parameter μ_i and σ_i^2 , we also obtain the same iteration as the Hasselblad method. Therefore, we can obtain the iterations naturally by maximizing the Q -function in the EM algorithm.

It is proved that the log likelihood increases in each step of the EM algorithm. Akamine (2007) used the geometrical method (Murata and Ikeda 2000) for this proof. For discrete models, the K-L information quantity is defined as

$$\text{KL}(\xi, \eta) = \sum_x \xi(x) \ln \frac{\xi(x)}{\eta(x)} \geq 0, \quad (3.38)$$

where $\xi(x)$ and $\eta(x)$ are probability distributions. This inequality is proved by using $\ln x \leq x - 1$ as follows (Sakamoto *et al.* 1986):

$$\sum_x \xi(x) \ln \frac{\eta(x)}{\xi(x)} \leq \sum_x \xi(x) \left(\frac{\eta(x)}{\xi(x)} - 1 \right) = \sum_x \eta(x) - \sum_x \xi(x) = 1 - 1 = 0. \quad (3.39)$$

The equality is clearly satisfied only when $\xi(x) \equiv \eta(x)$. The K-L information quantity gives the upper limit of the log likelihood as

$$Y = \sum_x H(x) \ln G(x) \leq \sum_x H(x) \ln H(x). \quad (3.40)$$

Define the distance as

$$D = \text{KL}(z_i(x), p_i g_i(x)) \geq 0. \quad (3.41)$$

This is a special distance because

$$\text{KL}(\xi, \eta) \neq \text{KL}(\eta, \xi). \quad (3.42)$$

Firstly, to minimize D we can modify $z_i(x)$ in E-step. We can divide D as follows:

$$\begin{aligned} D &= \sum_x \sum_i z_i(x) \ln \frac{z_i(x)}{p_i g_i(x)} \\ &= \sum_x \sum_i z_i(x) \ln \frac{z_i(x)}{H(x) p_i g_i(x) / G(x)} + \sum_x H(x) \ln \frac{H(x)}{G(x)} \\ &= \text{KL} \left(z_i(x), \frac{H(x) p_i g_i(x)}{G(x)} \right) + \text{KL} \left(H(x), G(x) \right). \end{aligned} \quad (3.43)$$

Therefore, the equations

$$z_i(x) = \frac{H(x) p_i g_i(x)}{G(x)} \quad (3.44)$$

make D minimum because the first term of Eq. (3.43) is 0 and the second term does not include $z_i(x)$.

Secondly, to minimize D we can modify the parameters in M-step. The distance D is divided in another way

$$D = \sum_x \sum_i z_i(x) \ln z_i(x) - \sum_x \sum_i z_i(x) \ln(p_i g_i(x)), \quad (3.45)$$

Because the first term is constant, we have to minimize the second term which is the same as the minus Q -function (3.32). Thus, it was proved that the distance D is made shorter in each E- and M-step. In E-step, D is modified as

$$D = \text{KL}(H(x), G(x)) = \sum_x H(x) \ln H(x) - Y. \quad (3.46)$$

In M-step, D is minimized. Thus, the objective function Y is increased in each step. It has an upper limit as Eq. (3.40). Then it must be converged. Although the convergence of the objective function Y is proved, the convergence of each parameter has not yet been proved.

3-5. Convergence criterion by diminishing mapping

On the estimation of the mixture ratio, Akamine (2001) showed the following convergence criterion according to Iri (1981). Iteration (3.17) can be rewritten as

$$p_i^{(t+1)} = y_i \left(p_1^{(t)}, \dots, p_n^{(t)} \right). \quad (3.47)$$

We obtain the following inequality:

$$\begin{aligned} \sum_i |dp_i^{(t+1)}| &= \sum_i \left| \sum_k \frac{\partial y_i}{\partial p_k} dp_k \right|^{(t)} \leq \sum_i \sum_k \left| \frac{\partial y_i}{\partial p_k} \right|^{(t)} |dp_k^{(t)}| \\ &\leq \max \left(\sum_i \left| \frac{\partial y_i}{\partial p_k} \right|^{(t)} \right) \sum_k |dp_k^{(t)}| = |\mathbf{J}^{(t)}| \sum_i |dp_i^{(t)}|, \end{aligned} \quad (3.48)$$

where $|\mathbf{J}^{(t)}|$ is the norm of the matrix $\mathbf{J}^{(t)}$. The elements of this are

$$\frac{\partial y_i}{\partial p_k} = - \sum_x \frac{H p_i g_i g_k}{G^2}, \quad (i \neq k) \quad (3.49)$$

$$\frac{\partial y_k}{\partial p_k} = - \sum_x \frac{H g_k (G - p_k g_k)}{G^2} = \sum_x \sum_{i \neq k} \frac{H g_k p_i g_i}{G^2} = \Psi_k, \quad (3.50)$$

where the \sum with $i \neq k$ means the sum of $i = 1, \dots, n$ except $i = k$. Thus, we obtain the following criterion:

$$|\mathbf{J}^{(t)}| = \max \left(\sum_i \left| \frac{\partial y_i}{\partial p_k} \right|^{(t)} \right) = 2 \max \Psi_k^{(t)} \leq L < 1, \quad (3.51)$$

where L is the Lipschitz constant. This is the sufficient condition of the convergence. On the other hand, the necessary condition is

$$\frac{\partial y_k}{\partial p_k} = \Psi_k \leq 1 \quad (k = 1, \dots, n). \quad (3.52)$$

3-6. Approximation of the Jacobi method

Although the Hasselblad method is regarded as the EM algorithm, as detailed above, it is also regarded as an approximation of the Jacobi method in which all parameters are modified by using

the Newton method at the same time in each iteration (Akamine 1998, 1999). Let Eq. (3.28) be

$$f_i = \sum_x \frac{H g_i}{G} - 1 = 0. \quad (3.53)$$

Thus, we obtain the following equation:

$$\frac{\partial f_i}{\partial p_i} = - \sum_x \frac{H g_i^2}{G^2} = - \frac{1}{p_i} \left[\sum_x \frac{H g_i}{G} - \sum_x \frac{H g_i (G - p_i g_i)}{G^2} \right]. \quad (3.54)$$

In the neighborhood of the solution, we obtain the approximation

$$\frac{\partial f_i}{\partial p_i} \sim - \frac{1 - \Psi_i}{p_i}, \quad (3.55)$$

where Ψ_i is Eq. (3.50). Therefore, the modification of p_i in the Jacobi method is

$$\Delta p_i = - \left(\frac{f_i}{\frac{\partial f_i}{\partial p_i}} \right)^{(t)} \sim \left[\frac{p_i \left(\sum_x \frac{H g_i}{G} - 1 \right)}{1 - \Psi_i} \right]^{(t)}. \quad (3.56)$$

This shows that the Hasselblad method is regarded as an approximation of the Jacobi method when Ψ_i is small.

3-7. Difference between the iteration method and the EM algorithm

There is little difference between the iteration method and the EM algorithm (Akamine 2005). It is clear when $\sigma_i = \alpha \mu_i$. In this case of the EM algorithm, the condition of the solution

$$\frac{\partial Q}{\partial \alpha^2} = 0 \quad (3.57)$$

gives the equation

$$\sum_x \sum_i z_i(x) \left\{ \alpha^2 - \frac{(x - \mu_i)^2}{\mu_i^2} \right\} = 0. \quad (3.58)$$

Therefore, we obtain the iteration

$$\alpha^{2(t+1)} = \sum_x \sum_i z_i(x)^{(t)} \frac{(x - \mu_i^{(t+1)})^2}{\mu_i^{(t+1)2}}. \quad (3.59)$$

On the other hand, the condition of the solution

$$\frac{\partial Q}{\partial \mu_i} = 0 \quad (3.60)$$

gives the quadratic equation

$$\sum_x z_i(x) \{ x(x - \mu_i) - \alpha^2 \mu_i^2 \} = 0. \quad (3.61)$$

The iteration is

$$\begin{aligned} \mu_i^{(t+1)} &= \frac{-B + \sqrt{B^2 + 4AC}}{2A} > 0, \\ A &= \alpha^{2(t+1)} \sum_x z_i(x)^{(t)} > 0, \\ B &= \sum_x z_i(x)^{(t)} x, \\ C &= \sum_x z_i(x)^{(t)} x^2 > 0. \end{aligned} \quad (3.62)$$

Equations (3.59) and (3.62) are impossible to solve because the values of α and μ on the right side are new values. In the iteration method, we can solve these equations by making the values of α or μ on the right side be old values, which is equal to the Hasselblad method. Therefore, in this case, M-step of the EM algorithm needs the iteration method. On the other hand, Akamine (2005) presented the Newton method for Eq. (3.61) as

$$\mu_i^{(t+1)} = \frac{A\mu_i^{(t)2} + C}{2A\mu_i^{(t)} + B}. \quad (3.63)$$

This method converges in the same way as the Hasselblad method.

(Example 3) Estimation of the age composition for the data of the porgy.

The body-length composition is shown in Table 3 and Fig. 6. Let the histogram data be $H^*(7.5) = 7$, $H^*(8.5) = 79$, and so on. The iterations (3.17), (3.20) and (3.23) can be calculated by using the spread-sheet software (Aizawa and Takiguchi 1999; Gorie 2002). Table 4 shows the results.

3-8. Marquardt method

The Newton method is a basic method for the optimization of the multivariable model

$$f_i(\theta_1, \dots, \theta_n) = 0, \quad f_i = \frac{\partial Y}{\partial \theta_i}, \quad (3.64)$$

Table 3. The body-length composition data of the porgy (*Dentex tumifrons*). BL is the body length (cm). Data were excerpted with kind permission of Dr. Shoichi Tanaka from Table 2 in Tanaka (1956) with its original data source of Group of Statistics, Pelagic Resource Section, SEIKAI Regional Fisheries Research Laboratory, Data of the investigation of September to November in 1950, Report of the Investigations of Trawl Fishery in the China Sea, No. 2, pp. 1–28, 1951. SEIKAI Regional Fisheries Research Laboratory, Nagasaki, Japan.

BL	Frequency	BL	Frequency	BL	Frequency
7–8	7	17–18	448	27–28	114
–9	79	–19	512	–29	64
–10	509	–20	719	–30	22
–11	2240	–21	673	–31	0
–12	2341	–22	445	–32	2
–13	623	–23	341	–33	2
–14	476	–24	310	–34	0
–15	1230	–25	228	–35	0
–16	1439	–26	168	–36	1
–17	921	–27	140	Total	14054

Table 4. The result of the estimation for the mixture of normal distributions. The range of the body length is 7–32 cm. Data were excerpted with permission from Akamine T. *Suisan Shigen Kaiseki no Kiso*, Table 1.2, Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku.

	1	2	3	4	5	m	$\ln L$
p	0.4106	0.3042	0.1916	0.0603	0.0332	14051	–37627.45
μ	11.00	15.28	19.79	23.46	26.75		
σ	0.873	1.133	1.480	1.265	1.423		

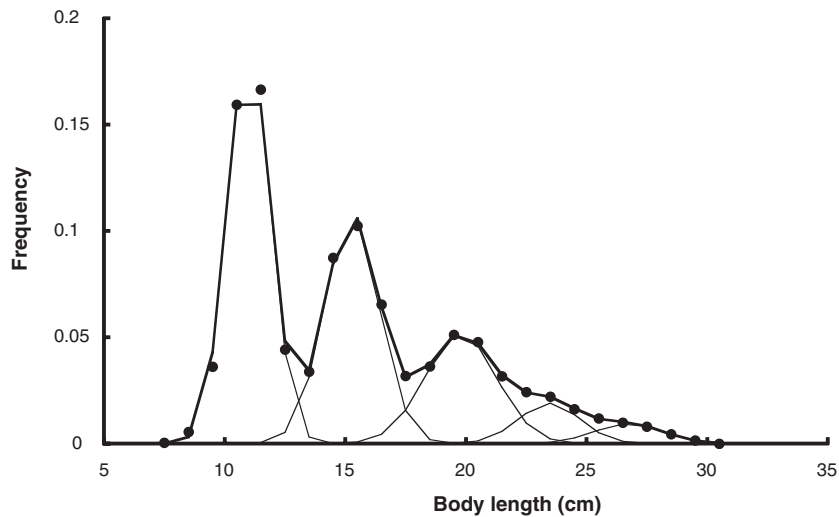


Fig. 6. Result of the estimation (Table 4). Black circles are data, thick line is a mixture, and thin lines are normal distributions. Redrawn with permission after Akamine T. *Suisan Shigen Kaiseki no Kiso*, p. 6, Fig. 1.3, Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku Co., Ltd.

where Y is the objective function and θ_i is a parameter. These are non-linear equations. The Newton method to solve these equations becomes an iteration method by using the following linear equations:

$$\mathbf{H}\Delta\theta = -\mathbf{f},$$

$$\mathbf{H} = \left(\frac{\partial f_i}{\partial \theta_j} \right) = \left(\frac{\partial^2 Y}{\partial \theta_i \partial \theta_j} \right), \quad \mathbf{f} = \left(\frac{\partial Y}{\partial \theta_i} \right), \quad (3.65)$$

where \mathbf{H} is the n -th square matrix called a Hessian, $\Delta\theta$ is the modified vector of parameters, and $-\mathbf{f}$ is the steepest descent vector. Although the convergence area of this is narrow, the Marquardt (1963) method was invented to enlarge the convergence area as follows:

$$(\mathbf{H} + \lambda \mathbf{I})\Delta\theta = -\mathbf{f}, \quad (3.66)$$

where \mathbf{I} is the unit matrix, and λ is an adjusting factor. When λ is large, Eq. (3.66) approaches to

$$\Delta\theta = -\frac{1}{\lambda} \mathbf{f}. \quad (3.67)$$

This is the steepest descent method and the convergence area is not large. Therefore, we use the standard scaling of parameters (Marquardt 1963; Akamine 1988b) to enlarge the convergence area and obtain

$$(\mathbf{H} + \lambda \mathbf{D})\Delta\theta = -\mathbf{f}, \quad (3.68)$$

where

$$\mathbf{D} = \text{diag} \mathbf{H} = \begin{cases} \frac{\partial^2 Y}{\partial \theta_i^2} & (i = j), \\ 0 & (i \neq j). \end{cases}$$

Thus, only the diagonal elements of \mathbf{H} need to be multiplied by $(1 + \lambda)$ for the standard scaling of parameters. When λ is large, Eq. (3.68) approaches

$$\Delta\theta_i = -\frac{1}{\lambda} \frac{\partial Y / \partial \theta_i}{\partial^2 Y / \partial \theta_i^2} = -\frac{1}{\lambda} \frac{f_i}{\partial f_i / \partial \theta_i}. \quad (3.69)$$

This is the Jacobi method (Akamine 1987b). Then the Marquardt method is a mixture of the Newton and the Jacobi methods, not a mixture of the Newton and the steepest descent methods. In the first half of the iteration, λ must be large enough to approach the Jacobi method. In the second half, λ must be small enough to approach the Newton method.

The procedure of the Marquardt method is as follows:

- (1) Let the initial value be $\lambda^{(0)} = 1$, and solve Eq. (3.68).
- (2) When Y is decreased, let $\lambda^{(t+1)} = \lambda^{(t)}/2$, and repeat the procedure (1)–(2).
- (3) When Y is not decreased, let $\lambda^{(t)} \leftarrow 2\lambda^{(t)}$ and solve Eq. (3.68) again. If Y is not decreased after 10 times repeating the same procedure, either the solution has been obtained or the initial values of the parameters are not adequate.

Figure 7 shows the image of the convergence (Akamine 1987a). The disadvantage point of this method is the calculation amount of the Hessian is large.

In numerical calculation, condition (3.2) is treated by using the following methods: One way is the elimination of p_n as

$$p_n = 1 - \sum_{i=1}^{n-1} p_i, \quad (3.70)$$

another way is the transformation of p_i to K_i as

$$p_i = \frac{K_i}{S}, \quad S = \sum_{i=1}^n K_i. \quad (3.71)$$

These ways change the conditional optimization to a normal optimization.

4. Estimation of the population size

The population size estimation is important for fisheries stock assessment. In this section, we discuss the basic methods for estimation of the population size which are based on random sampling. These are the Petersen method, the quadrat method, and the DeLury removal method. Their statistical models are based on the binomial distribution. The classical Bayesian statistics and sum formulae of the binomial and hyper-geometric distributions are also explained. The graphic function of spread-sheet software helps to understand these methods.

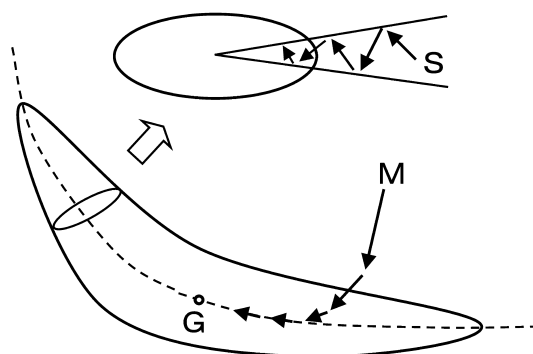


Fig. 7. Image of the convergence of the Marquardt method. M is the starting point, G is the goal, and S is the starting point of the steepest descent method which does not converge to the goal. Redrawn with permission after Akamine T. *Suisan Shigen Kaiseki no Kiso*, p. 11, Fig. 1.6, Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku Co., Ltd.

4-1. Petersen method

This is a basic method to estimate the number of fish by using the mark–recapture experiments. Let N be the total number of individuals to estimate, M be the number of marked in N , n be the total number of captured, and r be the number of mark–recaptures in n . Kuno (1986) showed the point estimation

$$N = \frac{Mn}{r}, \quad (4.1)$$

its variance

$$V(N) = \frac{Mn(M-r)(n-r)}{r^3}, \quad (4.2)$$

and its 95% confidence interval

$$N = \frac{Mn}{r} \pm 1.96\sqrt{V(N)}. \quad (4.3)$$

However, the precision of this interval is not high because it uses only propagation of errors.

The strict statistical model for the Petersen method is the hyper-geometric distribution

$$\text{HG}(r, n, M, N) = \frac{\binom{M}{r} \binom{N-M}{n-r}}{\binom{N}{n}}, \quad (4.4)$$

where

$$\binom{y}{x} = \frac{y!}{x!(y-x)!} = \frac{y^{(x)}}{x!} = \frac{y(y-1)\cdots(y-x+1)}{x(x-1)\cdots 1}. \quad (4.5)$$

When N and M is larger than n and r greatly, this approximates to the binomial distribution as follows:

$$\text{HG}(r, n, M, N) = \binom{n}{r} \frac{M^{(r)}(N-M)^{(n-r)}}{N^{(n)}} \sim \binom{n}{r} p^r (1-p)^{n-r} = \text{Bi}(r, n, p), \quad (4.6)$$

where $p = M/N$ is the mark ratio.

If p in the binomial distribution model is estimated, then we can obtain the total number of individuals as $N = M/p$. In practice, the approximation to the normal distribution easily gives the interval estimation. According to the normal distribution, the following equation gives the confidence interval of p :

$$z = \frac{r - np}{\sqrt{np(1-p)}}, \quad (4.7)$$

where z is the standardized variable. This is rewritten as

$$z^2 np(1-p) = (r - np)^2. \quad (4.8)$$

This is a quadratic equation of p . Akamine (2002) showed the roots of this equation with the continuity correction are

$$p = \frac{1}{n + z^2} \left[r \pm 0.5 + \frac{z^2}{2} \mp z \sqrt{(r \pm 0.5) \left(1 - \frac{r \pm 0.5}{n} \right) + \frac{z^2}{4}} \right]. \quad (4.9)$$

The reason that the sign in front of z is inversed is explained below. The roots without the continuity correction are

$$p = \frac{1}{n + z^2} \left[r + \frac{z^2}{2} \pm z \sqrt{r \left(1 - \frac{r}{n} \right) + \frac{z^2}{4}} \right]. \quad (4.10)$$

This equation is shown in Yamada and Kitada (1997). When n is large, these are approximated to

$$p = \frac{1}{n} \left[r \pm z \sqrt{r \left(1 - \frac{r}{n} \right)} \right] = \frac{r}{n} \pm z \sqrt{V(p)}. \quad (4.11)$$

This equation is shown in many statistics textbooks.

(Example 4) Estimation of the 95% interval of N when $M = 60$, $n = 141$, and $r = 11$.

When $z = 1.96$, we can obtain the 95% confidence interval. The solutions resulting from Eq. (4.11) are $N = 490.67$ and 1778.01 , thus the 95% confidence interval of N is $[491, 1778]$. The solutions resulting from Eq. (4.10) are $N = 446.78$ and 1360.00 , thus that of N is $[447, 1360]$. The solutions resulting from Eq. (4.9) are $N = 461.63$ and 1284.13 , thus that of N is $[462, 1284]$. The higher the precision becomes, the shorter the interval becomes. The true value of this experiment is $N = 1200$ and $p = 0.05$ (Yamada and Kitada 1997). Let $r = 11.5$ with the continuity correction, then we obtain $z = 1.72 < 1.96$. Thus, this experiment is successful and the confidence interval must include the true value. The solution from Eq. (4.9) is the best because it includes the true value and it is the shortest. When the approximation error of the binomial distribution to the normal distribution is not small, the solution from Eq. (4.10) may be safer than Eq. (4.9). On the other hand, the solutions from Eq. (4.3) are $N = 374.70$ and 1163.48 , thus the 95% confidence interval of N is $[375, 1163]$. This interval does not include the true value.

The traditional interpretation of the confidence interval is as follows: The 95% confidence interval includes the true value about 95 times in 100 trials. Figure 8 shows the 95% confidence interval of the Petersen method. This is called the 95% confidence belt. Let the true value p_0 be the dotted line. The data r_0 is included in the 95% probability interval (the dotted line in the 95% confidence belt), and its confidence interval of p (the line segment at $r = r_0$) includes the true value p_0 . On the other hand, the data r_1 is not included in the 95% probability interval, and its confidence interval (the line segment at $r = r_1$) does not include the true value p_0 . For interval estimation, the probability for horizontal (r) is only used, and the probability for vertical (p) is not used in traditional statistics.

The reason that the sign in front of z in Eq. (4.9) is inversed is as follows: Figure 8 shows the two curves of Eq. (4.9). The right curve gives the upper limit of r and the left one gives the lower limit. For the continuity correction, we plus 0.5 to r in the upper limit, and minus 0.5 in the lower limit. Figure 8 shows that the lower limit of p uses the upper limit of r , and the upper limit of p uses the lower limit of r . Thus, the sign in front of z in Eq. (4.9) is inversed. On the other hand, Takeuchi and Fujino (1981) used the equation in which the sign in front of z in Eq. (4.9) is not inversed, because they gave rejection regions, not a confidence interval.

4-2. Bayesian statistical method for the Petersen method

The classical Bayesian statistics method is easily applied for this model. Bayesian statistics is based on the following theorem: The posterior distribution of θ is

$$P(\theta|X) = \frac{L(\theta|X)\pi(\theta)}{\int L(\theta|X)\pi(\theta)d\theta}, \quad (4.12)$$

where X is data, θ is parameter, $L(\theta|X)$ is likelihood, and $\pi(\theta)$ is the prior distribution of θ . This equation means the following proportional expression:

$$P(\theta|X) \propto L(\theta|X)\pi(\theta). \quad (4.13)$$

All methods which treat parameter probability are involved in Bayesian statistics. Equation (4.12) is also regarded as the weighted likelihood.

HDR (Highest Density Region) can be defined as follows: Let

$$P(x) \geq P(y), \quad (4.14)$$

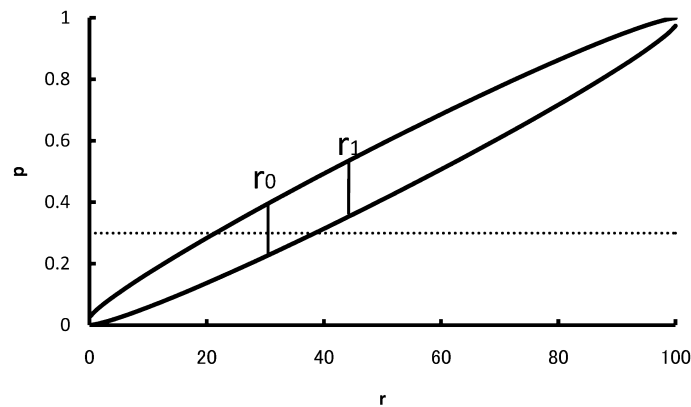


Fig. 8. The 95% confidence belt of the Petersen method. Redrawn with permission after Akamine T. *Suisan Shigen Kaiseki no Kiso*, p. 54, Fig. 3.5, Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku Co., Ltd.

where P is probability, x is the inner point of HDR, and y is the outside point of HDR. Thus, the area of HDR is the minimum. It is not always a continuous area for the mixed distribution. The following function can be defined as

$$\delta(P, h) = \begin{cases} 1, & P(x) \geq h, \\ 0, & P(x) < h. \end{cases} \quad (4.15)$$

This function gives the $(1 - \alpha)$ probability interval which is determined by h as

$$I(h) = \int P(x) \delta(P, h) dx = 1 - \alpha. \quad (4.16)$$

The favorable prior distribution of p is the uniform distribution from $p = 0$ to 1 for the Petersen method. This is the first example of Bayesian statistics in history. Let $\pi(p) = 1$. Thus, we obtain the posterior distribution

$$P(p|r) = (n+1) \text{Bi}(r, n, p), \quad (4.17)$$

because

$$\int_0^1 \text{Bi}(r, n, p) dp = \frac{1}{n+1}. \quad (4.18)$$

The following formula means that the upper probability of r and the lower probability of p are almost equal

$$(n+1) \int_0^p \text{Bi}(r, n, t) dt = p \text{Bi}(r, n, p) + \sum_{i=r+1}^n \text{Bi}(i, n, p). \quad (4.19)$$

When the binomial distribution approaches to the Poisson distribution not the normal distribution, the Bayesian result is not equal to the traditional one, because the Poisson distribution is not symmetric. When the rejection region of the normal distribution is 2.5% regions of both sides, and that of the Poisson distribution is an almost 5% region of the right side, so the results of the Bayesian method is different from the traditional one.

Solving Example 4 by using the Bayesian statistics, we obtain Fig. 9 as the posterior distribution of p with the prior distribution $\pi(p) = 1$ and the step-wise 0.005. If $h = 0.025$, we obtain the probability interval of p as [0.045, 0.130], and the total probability 95.32% > 95%. Removing $p = 0.130$ which has the minimum probability, we have the interval [0.045, 0.125] with the total probability 94.04% < 95%. In this case, the 95% probability interval of N is [480, 1333]. If $p = 0.130$ is included, the lower limit of N is 462, and becomes similar to the confidence interval of Eq. (4.9). When we use smaller steps, we obtain a higher precision interval.

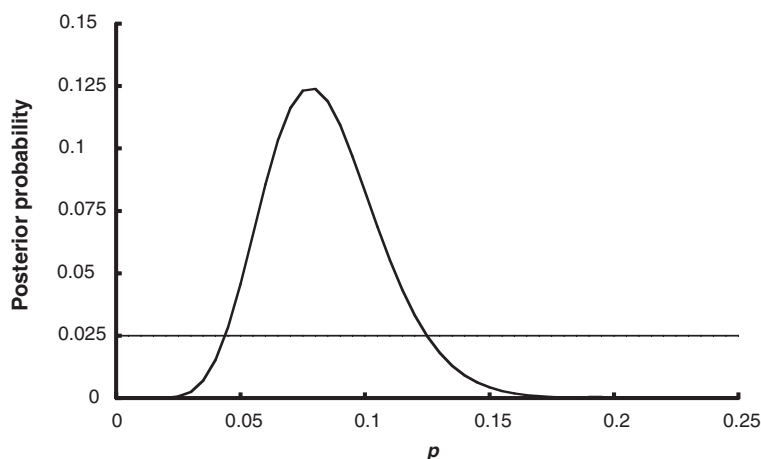


Fig. 9. The posterior distribution of the Petersen method. Redrawn with permission after Akamine T. *Suisan Shigen Kaiseki no Kiso*, p. 56, Fig. 3.6, Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku Co., Ltd.

4-3. Bayesian statistical method by using the hyper-geometric distribution

The strict statistical model for the Petersen method is the hyper-geometric distribution explained above. In the binomial distribution model, we use the uniform distribution for the prior distribution of p . Therefore, the prior distribution of N is derived as follows: Let $p = f(N)$. Thus,

$$\Delta p = f(N+1) - f(N) = \frac{M}{N+1} - \frac{M}{N} = -\frac{M}{(N+1)N}. \quad (4.20)$$

Then the favorable prior distribution of N is

$$\pi(N) = \frac{M+1}{(N+2)(N+1)}, \quad (4.21)$$

because the following equation holds:

$$\sum_{N=M+n-r}^{\infty} \frac{M+1}{(N+2)(N+1)} \text{HG}(r, n, M, N) = \frac{1}{n+1}. \quad (4.22)$$

Therefore, the posterior distribution of N is

$$P(N|r) = \frac{(n+1)(M+1)}{(N+2)(N+1)} \text{HG}(r, n, M, N). \quad (4.23)$$

The following equation means that the upper probability of r and the upper probability of N are almost equal

$$\begin{aligned} & \sum_{i=N}^{\infty} \frac{(n+1)(M+1)}{(i+2)(i+1)} \text{HG}(r, n, M, i) \\ &= \frac{M+1}{N+1} \text{HG}(r, n, M, N) + \sum_{i=r+1}^n \text{HG}(i, n, M+1, N+1). \end{aligned} \quad (4.24)$$

This model is just a simple model in which the binomial distribution is transformed to the hyper-geometric distribution. Thus, $(n+1)\text{HG}(r, n, M, N)$ must be the probability density in the posterior distribution. When the prior distribution is not a uniform distribution, the posterior probability is not equal to the likelihood. We had better make the probability interval of the posterior distribution be almost equal to the confidence interval of the traditional statistics (Akamine 1989b).

4.4. Quadrat method

This is a basic method to estimate the total number of fish, plankton, or benthos by counting the number of individuals in a part (Akamine 1981, 2002). When fish distribute at random, the number of individuals caught by a part is modeled as the binomial distribution:

$$\text{Bi}(r, n, p) = \binom{n}{r} p^r (1-p)^{n-r}, \quad (4.25)$$

where n is the total number, p is the fishing ratio, and r is the catch. This is the statistical model to estimate n by using p and r . Although the point estimation is easy as $n = r/p$, the interval estimation is not so easy. We obtained the following estimation:

$$n = \frac{r}{p} \pm z \sqrt{V(n)}, \quad (4.26)$$

$$V(n) = \frac{V(r)}{p^2} = \frac{np(1-p)}{p^2} = \frac{r(1-p)}{p^2}.$$

This is similar to Eq. (4.3), so the precision is not high.

In practice, the approximation to the normal distribution gives the interval estimation. Equation (4.8) is also a quadratic equation of n . The roots of Eq. (4.8) with the continuity correction are

$$n = \frac{1}{p} \left[r \pm 0.5 + \frac{z^2}{2}(1-p) \mp z \sqrt{(r \pm 0.5)(1-p) + \frac{z^2}{4}(1-p)^2} \right]. \quad (4.27)$$

The reason that the sign in front of z is inverted is as same as for the Petersen method. **Figure 10** shows two curves of these, they are called the 95% confidence belt, when $z = 1.96$ and $p = 0.2$.

4.5. Bayesian statistical method for the quadrat method

The favorable prior distribution of n is the uniform distribution from $n = r$ to ∞ for the quadrat method. Let the prior distribution be $\pi(n) = \epsilon > 0$ (Mangel and Beder 1985; Akamine 1989a). Thus, we obtain the posterior distribution

$$P(n|r) = p\text{Bi}(r, n, p), \quad (4.28)$$

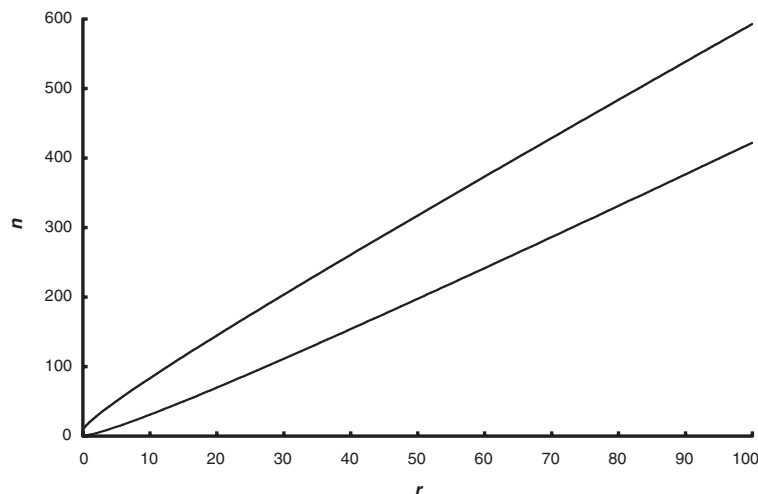


Fig. 10. The 95% confidence belt of the quadrat method. Redrawn with permission after Akamine T. *Suisan Shigen Kaiseki no Kiso*, p. 51, Fig. 3.3, Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku Co., Ltd.

because

$$\sum_{n=r}^{\infty} \text{Bi}(r, n, p) = \frac{1}{p}. \quad (4.29)$$

(Example 5) Estimation of the 95% interval of n when $r = 5$ and $p = 0.1$.

The solutions of Eq. (4.27) are $n = 25.38, 105.35$. Thus, the 95% confidence interval of n is [26, 105].

On the other hand, the posterior distribution of n is shown in Fig. 11 by using the uniform distribution for the prior distribution in Bayesian statistics. This is the vertical probability of Fig. 10. When $h = 0.0025$, the interval of n is [19, 105]. The total probability of this case is 95.16%, which is larger than 95%. The total probability without $n = 105$ (the probability of which is minimum) is 94.91%, thus the 95% probability interval of n is [19, 104]. This interval is similar to the traditional confidence interval. The following formula supports this phenomenon.

$$\sum_{i=r}^n \text{Bi}(i, n, p) = p \sum_{j=r}^n \text{Bi}(r-1, j-1, p). \quad (4.30)$$

This formula suggests that the upper probability of r and the lower probability of n are almost equal. When the binomial distribution approaches to the Poisson distribution, the Bayesian estimate and the traditional one are not equal.

4-6. DeLury removal method

This method gives the estimation of the initial population size N_0 and the catchability coefficient q at the same time. This method is modeled in two ways as a linear regression and a product of binomial distributions. Although Kuno (1986) recommended the former in practice, Akamine *et al.* (1992) recommended the expanded least-squares method based on the latter. Their objective function is given by using the approximation to the normal distribution as follows:

$$Y = \sum_{t=0}^n \frac{(C_t - N_t p_t)^2}{N_t p_t (1 - p_t)}, \quad (4.31)$$

$$p_t = q X_t, \quad N_t = N_0 - K_t, \quad K_t = \sum_{i=0}^{t-1} C_i,$$

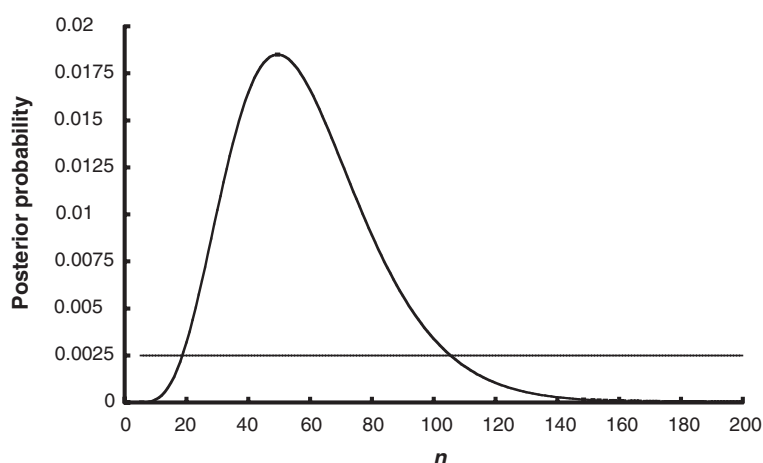


Fig. 11. The posterior distribution of the quadrat method. Redrawn with permission after Akamine T. *Suisan Shigen Kaiseki no Kiso*, p. 52, Fig. 3.4, Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku Co., Ltd.

where C_t is the number of fish caught, N_t is the population size, p_t is the removal ratio, and X_t is the fishing effort at t -th time. This is the expanded least-squares method. Therefore, the distribution of the minimum value of Y with respect to N_0 and q is $\chi^2(n-1)$. Although this equation was deduced by Zippin (1963) for the minimum χ^2 method, Akamine *et al.* (1992) obtained it from the likelihood ratio test. However, the confidence region of this model is very wide.

On the other hand, the product of binomial distributions is modified to the multinomial distribution. Thus, the other objective function is given as

$$Y^* = \sum_{t=0}^n \frac{(C_t - N_0 s_t)^2}{N_0 s_t} + \frac{(N_0 - K_{n+1} - N_0 v)^2}{N_0 v}, \quad (4.32)$$

$$s_t = p_t \prod_{i=0}^{t-1} (1 - p_i), \quad v = \prod_{i=0}^n (1 - p_i).$$

The distribution of the minimum value of this with respect to N_0 and q is also $\chi^2(n-1)$. Nowadays, the confidence region of N_0 and q with $\chi^2(2)$ is given by computer software.

(Example 6) Analysis of the data in Table 5.

The optimization method of spread-sheet software gives the minimum value of Eq. (4.31) as $N_0 = 1371.4$, $q = 0.0010651$, and $Y_{\min} = 10.428$. This minimum value of Y is smaller than $\chi^2(13) = 22.36$ (5%). Thus, this data passes the test of fitness. We can obtain the 95% confidence region by using the optimization method with the condition $Y_0 = Y_{\min} + \chi^2(2) = 16.419$, where $\chi^2(2) = 5.991$ (5%). This contour of Y_0 is shown in Fig. 12. The maximum N_0 is 2171.9 ($q = 0.00060$), and minimum is 1074.9 ($q = 0.00150$). Therefore, the confidence interval of N_0 is [1075, 2171]. It is important that the upper range of the confidence interval is larger than the lower one.

4-7. Proof of the sum formulae of the binomial distribution and the hyper-geometric distribution

Akamine (2002) proved some formulae as follows: The basic formula is

$$\text{Bi}(r, n, p) = p\text{Bi}(r-1, n-1, p) + (1-p)\text{Bi}(r, n-1, p). \quad (4.33)$$

The sum of both sides is

$$\sum_{i=r}^n \text{Bi}(i, n, p) = p\text{Bi}(r-1, n-1, p) + \sum_{i=r}^{n-1} \text{Bi}(i, n-1, p). \quad (4.34)$$

Table 5. Data for the DeLury removal method. Data were excerpted with permission from *Nippon Suisan Gakkaishi*, 24, Nose Y. On the confidence limits corresponding to the estimate obtained by DeLury's logarithmic catch-effort method. 953–956, 1959. ©1959, the Japanese Society of Fisheries Science.

No.	Effort ($\times 10^3$)	Catch	No.	Effort ($\times 10^3$)	Catch
1	20	28	9	50	41
2	42	70	10	34	28
3	86	121	11	24	23
4	67	77	12	26	28
5	88	95	13	18	14
6	66	59	14	9	5
7	68	71	15	4	2
8	31	22	Total	633	684

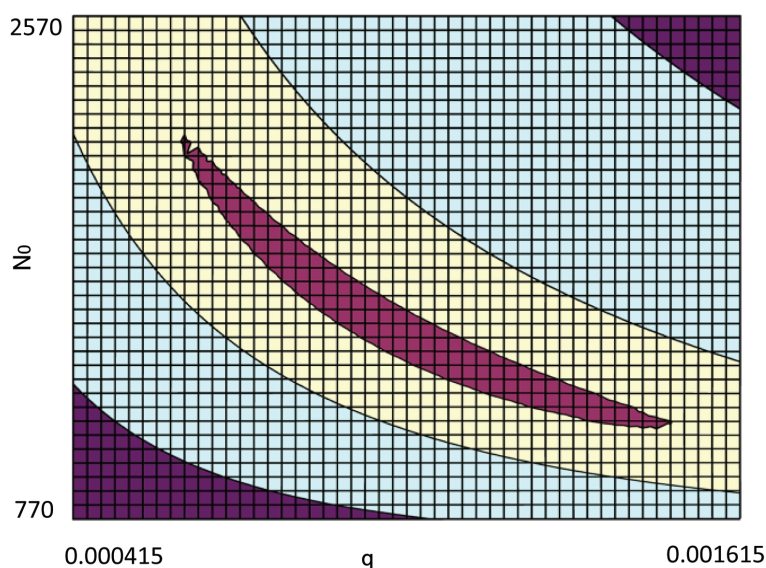


Fig. 12. The 95% confidence area of the DeLury removal method. Redrawn with permission after Akamine T. *Suisan Shigen Kaiseki no Kiso*, p. 61, Fig. 3.8, Kouseisha Kouseikaku, Tokyo, 2007. ©2007, Kouseisha Kouseikaku Co., Ltd.

The recursion of this gives Eq. (4.30). The limit is Eq. (4.29). On the other hand, integral by parts gives

$$\begin{aligned}
 \int_0^p \text{Bi}(r, n, t) dt &= \binom{n}{r} \int_0^p t^r (1-t)^{n-r} dt \\
 &= \binom{n}{r} \left\{ \left[\frac{t^{r+1}}{r+1} (1-t)^{n-r} \right]_0^p + \frac{n-r}{r+1} \int_0^p t^{r+1} (1-t)^{n-r-1} dt \right\} \\
 &= \frac{1}{n+1} \text{Bi}(r+1, n+1, p) + \int_0^p \text{Bi}(r+1, n, t) dt \\
 &= \frac{1}{n+1} \sum_{i=r}^n \text{Bi}(i+1, n+1, p). \tag{4.35}
 \end{aligned}$$

The limit is Eq. (4.18), and Eqs. (4.34) and (4.35) give Eq. (4.19).

For the hyper-geometric distribution, as the same way as the binomial distribution, we obtain the following formulae:

$$\begin{aligned}
 \text{HG}(r, n, M, N) &= \frac{M-r+1}{N-n+1} \text{HG}(r-1, n-1, M, N) \\
 &\quad + \frac{N-M-n+r+1}{N-n+1} \text{HG}(r, n-1, M, N), \tag{4.36}
 \end{aligned}$$

$$\sum_{i=r}^n \text{HG}(i, n, M, N) = \frac{M}{N} \text{HG}(r-1, n-1, M-1, N-1) + \sum_{i=r}^{n-1} \text{HG}(i, n-1, M, N), \tag{4.37}$$

$$\sum_{i=r}^n \text{HG}(i, n, M, N) = \frac{M}{N} \sum_{j=r}^n \text{HG}(r-1, j-1, M-1, N-1), \tag{4.38}$$

and

$$\sum_{n=r}^{N-M+r} \text{HG}(r, n, M, N) = \frac{N+1}{M+1}. \quad (4.39)$$

Although many formulae of the hyper-geometric distribution can be derived by the use of the recurrence formula, Eqs. (4.22) and (4.24) are derived by the use of the summation by parts. The difference is defined as follows:

$$\Delta f(x) = f(x+1) - f(x). \quad (4.40)$$

Hence, we obtain

$$\sum_{x=a}^b \Delta f(x) = [f(x)]_a^{b+1}. \quad (4.41)$$

Let $\Delta f(x) = g(x)$, this gives the indefinite summation

$$\Delta^{-1} g(x) = f(x) + c, \quad (4.42)$$

where Δ^{-1} is the summation operator, c is the summation constant. Therefore, the definite summation is denoted as

$$\sum_{x=a}^b g(x) = [\Delta^{-1} g(x)]_a^{b+1}. \quad (4.43)$$

The difference of the product is

$$\Delta(f(x)g(x)) = \Delta f(x)g(x+1) + f(x)\Delta g(x). \quad (4.44)$$

Thus, we obtain the following two formulae of the summation by parts:

$$\sum_a^b f(x)g(x) = [\Delta^{-1} f(x)g(x-1)]_a^{b+1} - \sum_a^b \Delta^{-1} f(x)\Delta g(x-1), \quad (4.45)$$

$$\sum_a^b f(x)g(x) = [\Delta^{-1} f(x)g(x)]_a^{b+1} - \sum_a^b \Delta^{-1} f(x+1)\Delta g(x). \quad (4.46)$$

The differences of the combination with respect to x are

$$\Delta \binom{x}{r} = \binom{x}{r-1}. \quad (4.47)$$

and

$$\Delta \frac{1}{\binom{x}{r}} = -\frac{r}{r+1} \frac{1}{\binom{x+1}{r+1}}. \quad (4.48)$$

The inverse operation of this is

$$\Delta^{-1} \frac{1}{\binom{x}{r}} = -\frac{r}{r-1} \frac{1}{\binom{x-1}{r-1}} + c. \quad (4.49)$$

Formula (4.22) can be modified as

$$\sum_{N=M+n-r}^{\infty} \frac{(n+1)(M+1)}{(N+2)(N+1)} \text{HG}(r, n, M, N) = \frac{M+1}{n+2} \binom{M}{r} S = 1. \quad (4.50)$$

We use formula (4.45) for S .

$$\begin{aligned}
 S &= \sum_{N=M+n-r}^{\infty} \frac{\binom{N-M}{n-r}}{\binom{N+2}{n+2}} \\
 &= - \left[\frac{n+2}{n+1} \frac{\binom{N-M-1}{n-r}}{\binom{N+1}{n+1}} \right]_{M+n-r}^{\infty} + \sum \frac{n+2}{n+1} \frac{\binom{N-M-1}{n-r-1}}{\binom{N+1}{n+1}} \\
 &= - \left[\frac{n+2}{r+1} \frac{1}{\binom{N-n+r+1}{r+1}} \right]_{M+n-r}^{\infty} \\
 &= \frac{n+2}{(M+1) \binom{M}{r}}. \tag{4.51}
 \end{aligned}$$

Thus, formula (4.22) is proved. Next, the left side of formula (4.24) can be modified as

$$\sum_{i=N}^{\infty} \frac{(n+1)(M+1)}{(i+2)(i+1)} \text{HG}(r, n, M, i) = \frac{M+1}{n+2} \binom{M}{r} S^*. \tag{4.52}$$

We use formula (4.46) for S^* .

$$\begin{aligned}
 S^* &= \sum_{i=N}^{\infty} \frac{\binom{i-M}{n-r}}{\binom{i+2}{n+2}} = - \left[\frac{n+2}{n+1} \frac{\binom{i-M}{n-r}}{\binom{i+1}{n+1}} \right]_N^{\infty} + \sum \frac{n+2}{n+1} \frac{\binom{i-M}{n-r-1}}{\binom{i+2}{n+1}} \\
 &= \sum_{j=r}^n \frac{n+2}{j+1} \frac{\binom{N-M}{j-r}}{\binom{N+1}{j+1}}. \tag{4.53}
 \end{aligned}$$

Therefore, we obtain the following equation:

$$\sum_{i=N}^{\infty} \frac{(n+1)(M+1)}{(i+2)(i+1)} \text{HG}(r, n, M, i) = \frac{M+1}{N+1} \sum_{j=r}^n \text{HG}(r, j, M, N). \tag{4.54}$$

Finally, Eqs. (4.54), (4.37) and (4.38) give Eq. (4.24).

5. Survival models

The virtual population analysis (VPA) is a typical survival model in fish population dynamics (Megrey 1989). Most of them use mortality coefficients. On the other hand, the Leslie matrix model is a famous management model in agriculture. The VPA model using mortality rates is useful because it is related to the Leslie matrix model, and linear programming. In this section, these models are presented as simple styles, because simple ones are most useful for many applications.

5-1. VPA

The standard fishing models are

$$\frac{dN}{dt} = -ZN = -(F + M)N, \tag{5.1}$$

$$\frac{dC}{dt} = FN, \quad (5.2)$$

where $N(t)$ is the number of fish, $C(t)$ is the number of fish caught, Z is the total mortality coefficient, F is the fishing mortality coefficient, and M is the natural mortality coefficient. The resulting equation from these two equations is

$$\frac{dN}{dt} + MN = -\frac{dC}{dt}. \quad (5.3)$$

Aksland (1994) and Hiramatsu (1995) obtained the analytical solution of this as

$$N(i) = N(i+1) \exp\left(\int_i^{i+1} M(t)dt\right) + \int_i^{i+1} \exp\left(\int_i^t M(s)ds\right) \frac{dC(t)}{dt} dt. \quad (5.4)$$

On the other hand, let $U(t)$ be the number of natural mortality and we have

$$\frac{dU}{dt} = MN. \quad (5.5)$$

Hence, Eqs. (5.1), (5.2) and (5.5) give a simple relationship

$$dN + dU + dC = 0. \quad (5.6)$$

Equations (5.1) and (5.2) also give the relationship

$$\frac{dC}{dN} = -\frac{F}{Z}. \quad (5.7)$$

When F and M are constants, the integrals of Eqs. (5.1) and (5.2) are

$$N(t) = e^{-(F+M)t} N_0, \quad (5.8)$$

$$C(t) = \frac{F}{F+M} (1 - e^{-(F+M)t}) N_0, \quad (5.9)$$

where $N_0 = N(0)$. VPA is a simple estimating method of N from C by using these fishing equations.

The definite calculation of VPA is as follows: The constant value of M is known. The basic equations are

$$N_{i+1} = e^{-(F_i+M)} N_i, \quad (5.10)$$

$$C_i = \frac{F_i}{F_i + M} (1 - e^{-(F_i+M)}) N_i, \quad (5.11)$$

where $N_i = N(i)$, and

$$C_i = \Delta C(i) = C(i+1) - C(i). \quad (5.12)$$

Substituting Eq. (5.10) into Eq. (5.11) gives

$$C_i = \frac{F_i}{F_i + M} (e^{F_i+M} - 1) N_{i+1}. \quad (5.13)$$

At first, we solve this non-linear equation for F_i , and then we obtain N_i of Eq. (5.10). The recurrence of these procedures gives all N_i . This method needs the number of fish in the oldest class N_T which is called the terminal N . In practice, we use the terminal F instead of N . The

following Ishioka and Kishida (1985) iteration method for Eq. (5.13) has a high precision and converges at about three times recurrence. Equation (5.13) can be modified as follows:

$$N_i = e^M N_{i+1} + \Psi(F_i, M) C_i, \quad \Psi(F, M) = \frac{F + M}{1 - e^{-F-M}} \frac{1 - e^{-F}}{F}. \quad (5.14)$$

This is a special case of Eq. (5.4). The Ishioka and Kishida iteration method is

$$F_i^{(t+1)} = \ln \left[1 + \frac{C_i}{N_{i+1}} \Psi(F_i^{(t)}, M) e^{-M} \right]. \quad (5.15)$$

On the other hand, a few approximations are used for this non-linear equation. The famous Pope method is

$$\Psi(F, M) \sim e^{M/2}. \quad (5.16)$$

Thus, Eq. (5.14) is rewritten as

$$N_i = e^M N_{i+1} + e^{M/2} C_i. \quad (5.17)$$

This is an impulse fishing model at the middle point in the fishing period. The fishing equation of this model is

$$C_i = e^{-M/2} (1 - e^{-F_i}) N_i. \quad (5.18)$$

The following Nagai (1980) method has a higher precision than the Pope method. The Nagai method is

$$\Psi(F, M) \sim \frac{M}{1 - e^{-M}}. \quad (5.19)$$

This is also an impulse fishing model in which the fishing point is a little before that of the Pope method. This approximation is also regarded as

$$\varphi(F + M) \sim \varphi(F) \varphi(M), \quad (5.20)$$

where

$$\begin{aligned} \Psi(F, M) &= \frac{\varphi(F + M)}{\varphi(F)}, \quad \varphi(x) = \frac{x}{1 - e^{-x}} = \phi(-x), \\ \phi(x) &= \frac{x}{e^x - 1} = 1 - \frac{x}{2} + \frac{x^2}{12} - \frac{x^4}{720} + \frac{x^6}{30240} - \frac{x^8}{1209600} + \dots \end{aligned} \quad (5.21)$$

The last function is the generating function of Bernoulli numbers. The complete solution of Eq. (5.20) is only the exponential function.

Akamine (2001) defined the impulse fishing as follows: Let

$$F(t) = \begin{cases} \frac{F}{b-a} & (i+a \leq t < i+b), \\ 0 & (i \leq t < i+a, \quad i+b \leq t < i+1). \end{cases} \quad (5.22)$$

Hence, we have

$$C_i = e^{-Ma} \frac{F}{F + M(b-a)} (1 - e^{-F-M(b-a)}) N_i. \quad (5.23)$$

When $b \rightarrow a$, we obtain

$$F(t) = F \delta(t - i - a), \quad (5.24)$$

where δ is the Dirac delta function, and

$$C_i = e^{-Ma} (1 - e^{-F}) N_i, \quad (5.25)$$

Equation (5.18) is a special case of this.

5-2. VPA using mortality rates

The following simple model is used for whales:

$$N_{i+1} = (N_i - C_i)e^{-M}, \quad (5.26)$$

$$C_i = E_i N_i, \quad (5.27)$$

where E is the fishing mortality rate. This is an impulse fishing model at the first point in the fishing period. The model becomes the same as the Pope method by sliding a fishing time for a half period.

The discrete fishing models are

$$\Delta N(i) = N_{i+1} - N_i = -(E_i + D_i)N_i, \quad (5.28)$$

$$\Delta C(i) = C_i = E_i N_i, \quad (5.29)$$

where D is the natural mortality rate. Thus the fishing models using mortality rates instead of mortality coefficients are as follows (Akamine 1995a, d):

$$N_{i+1} = s_i N_i = (1 - E_i - D_i)N_i, \quad (5.30)$$

$$C_i = E_i N_i, \quad (5.31)$$

where s is the survival rate. Hence, forward-calculation is

$$N_{i+1} = (1 - D_i)N_i - C_i \quad (5.32)$$

and back-calculation is

$$N_i = \frac{N_{i+1} + C_i}{1 - D_i}. \quad (5.33)$$

This is rewritten as

$$N_i = v_i N_{i+1} + v_i C_i, \quad (5.34)$$

where $v_i = 1/(1 - D_i)$. This is equivalent to Eq. (5.4) or Eq. (5.14) essentially. When D is constant, we have

$$N_0 = vC_0 + v^2C_1 + \cdots + v^T C_{T-1} + v^T N_T. \quad (5.35)$$

The relationships of mortality rates and mortality coefficients are

$$s = e^{-Z}, \quad (5.36)$$

$$E = \frac{F}{F + M} (1 - e^{-F-M}), \quad (5.37)$$

$$D = \frac{M}{F + M} (1 - e^{-F-M}). \quad (5.38)$$

The inverse transformations are

$$F = -\frac{E}{E + D} \ln(1 - E - D), \quad (5.39)$$

$$M = -\frac{D}{E + D} \ln(1 - E - D). \quad (5.40)$$

In these transformations, the following relationships hold:

$$E + D = 1 - e^{-F-M}, \quad \frac{E}{D} = \frac{F}{M}. \quad (5.41)$$

These are shown in Fig. 13, in which solid and broken lines have a one-to-one correspondence.

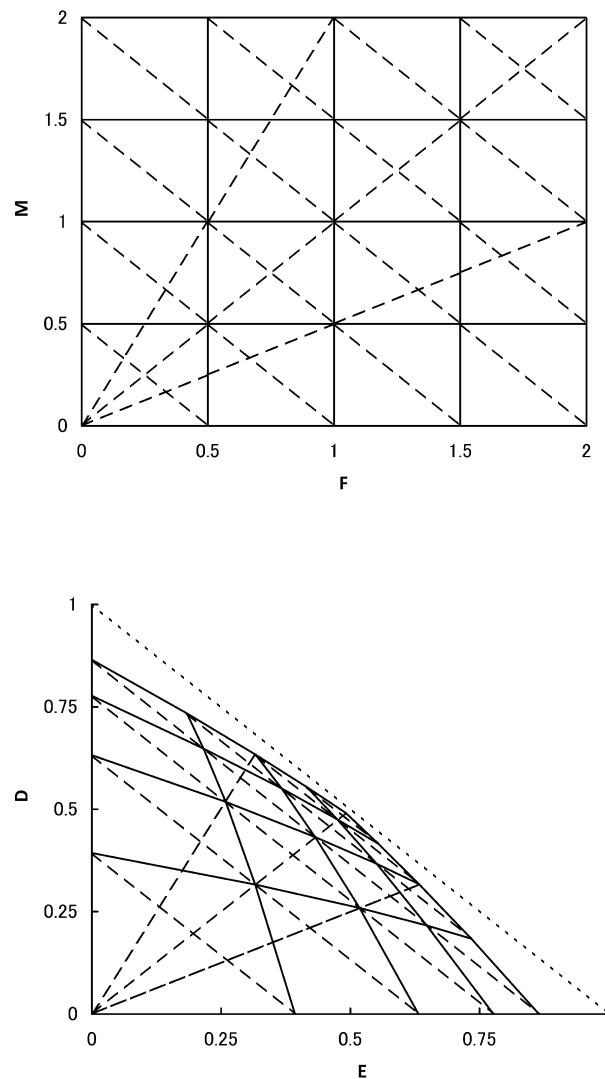


Fig. 13. The transformation of the coefficients and rates. Redrawn after *Bull. Jpn. Soc. Fish. Oceanogr.*, **59**, Fig. 1, Akamine T. Introduction to cohort analysis (VPA), 424–437, 1995, with permission from the Japanese Society of Fisheries Oceanography.

5-3. Leslie matrix model

This is a famous discrete survival model represented as

$$\mathbf{n}(j+1) = \mathbf{L}\mathbf{n}(j). \quad (5.42)$$

The representation by the components of this is

$$\begin{pmatrix} N_0(j+1) \\ N_1(j+1) \\ N_2(j+1) \\ \vdots \\ N_T(j+1) \end{pmatrix} = \begin{pmatrix} R_0 & R_1 & R_2 & \dots & R_T \\ s_0 & 0 & 0 & \dots & 0 \\ 0 & s_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & s_{T-1} & 0 \end{pmatrix} \begin{pmatrix} N_0(j) \\ N_1(j) \\ N_2(j) \\ \vdots \\ N_T(j) \end{pmatrix}, \quad (5.43)$$

where $N_i(j)$ is the number of i -th year old fish in j year, R_i is the reproduction rate and s_i is the survival rate of i -th year old fish. The resolution of this Leslie matrix is

$$\mathbf{L} = \mathbf{RS}, \quad (5.44)$$

$$\mathbf{R} = \begin{pmatrix} r_0 & r_1 & r_2 & \dots & r_T \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} s_0 & 0 & \dots & 0 \\ 0 & s_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & s_T \end{pmatrix},$$

where r_i is the number of eggs reproduced by one individual and $R_i = r_i s_i$. Thus, the forward-calculation of VPA is represented by matrix as follows (Akamine 1995b):

$$\mathbf{n}(j+1) = \mathbf{R}[(\mathbf{I} - \mathbf{D})\mathbf{n}(j) - \mathbf{c}(j)], \quad (5.45)$$

$$\mathbf{D} = \begin{pmatrix} D_0 & 0 & \dots & 0 \\ 0 & D_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & D_T \end{pmatrix}, \quad \mathbf{c}(j) = \begin{pmatrix} C_0(j) \\ C_1(j) \\ \vdots \\ C_T(j) \end{pmatrix}.$$

5-4. Linear programming for fishing equations

Akamine (1996, 1997) showed the optimization by linear programming for fishing. He rewrote Eq. (5.32) as

$$N_{i+1} = m_i N_i - C_i, \quad (5.46)$$

where m_i is constant. Resulting from Eq. (5.30) is

$$N_i = N_0 \prod_{i=0}^{n-1} s_i. \quad (5.47)$$

Definition of the reproduction is

$$R = r_1 N_1 + r_2 N_2 + \dots + r_n N_n \geq R_0 \quad (5.48)$$

and that of the yield is

$$Y = w_0 C_0 + w_1 C_1 + \dots + w_{n-1} C_{n-1}. \quad (5.49)$$

Let the fishing mortality rates be control variables as

$$0 \leq E_i \leq E_i^{\max}. \quad (5.50)$$

Then we obtain the inequalities

$$(m_i - E_i^{\max}) N_i \leq N_{i+1} \leq m_i N_i. \quad (5.51)$$

The objective function (5.49) is rewritten as

$$Y = w_0 m_0 N_0 + \sum_{i=1}^{n-1} (w_i m_i - w_{i-1}) N_i - w_{n-1} N_n. \quad (5.52)$$

Therefore, we can use the linear programming for the objective function (5.52) under conditions (5.48) and (5.51).

6. Summary

In this monograph, the author has presented four subjects. The first is the new standard growth formula and its statistical methods for estimation and test. The second subject is the body-size composition analysis in which we use the maximum likelihood method for parameter estimation on a mixture of normal distributions. The third subject is the estimation of population size in which the Petersen method, the quadrat method, and the DeLury removal method are discussed. The last subject is the survival models in which VPA, the Leslie matrix model, and the linear programming are explained. We can use all the methods in this monograph by using spread-sheet software.

In Section 2, a new standard growth formula based on the Richards growth formula and a periodic function

$$w(t) = \frac{w_{\infty}}{(1 + re^{-k[F(t)-F(c)]})^{1/r}}, \quad F(t) = t + \frac{A}{2\pi} \sin 2\pi(t - t_1)$$

is presented. Parameter estimation and test for this formula is based on the weighted least-squares method. The Ohnishi and Akamine growth formula, which is an implicit function, and the Awaysa method to estimate parameters of implicit models are introduced. The generalized reproduction model which is similar to the Richards growth formula is also explained.

In Section 3, the Hasselblad method for estimation on a mixture of normal distributions is discussed. This algorithm is obtained by using the EM algorithm naturally, and the convergence of its objective function is proved by using the K-L information quantity. The sufficient condition for convergence of parameters is as follows:

$$2 \max \Psi_k \leq L < 1, \quad \Psi_k = \sum_x \frac{H g_k (G - p_k g_k)}{G^2},$$

where L is the Lipschitz constant. On the other hand, the modification of the Jacobi method is

$$\Delta p_i = - \left(\frac{f_i}{\frac{\partial f_i}{\partial p_i}} \right) \sim \left[\frac{p_i \left(\sum_x \frac{H g_i}{G} - 1 \right)}{1 - \Psi_i} \right].$$

This shows the convergence speed of the Hasselblad method, which is the numerator of the right side, is similar to the Jacobi method.

In Section 4, the traditional methods in which the binomial distribution is approximated to the normal distribution for the Petersen method and the quadrat method are explained. The Bayesian statistical methods for these are also presented. The hyper-geometric distribution is used for the Petersen method in Bayesian, and its sum formulae are proved by using the formulae of summation by parts. The following formulae show the relationship between the traditional method and the Bayesian method:

$$(n+1) \int_0^p \text{Bi}(r, n, t) dt = p \text{Bi}(r, n, p) + \sum_{i=r+1}^n \text{Bi}(i, n, p),$$

$$\begin{aligned} & \sum_{i=N}^{\infty} \frac{(n+1)(M+1)}{(i+2)(i+1)} \text{HG}(r, n, M, i) \\ &= \frac{M+1}{N+1} \text{HG}(r, n, M, N) + \sum_{i=r+1}^n \text{HG}(i, n, M+1, N+1), \end{aligned}$$

and

$$\sum_{i=r}^n \text{Bi}(i, n, p) = p \sum_{j=r}^n \text{Bi}(j-1, n-1, p).$$

The DeLury removal method whose objective function is defined as the expanded weighted least-squares

$$Y = \sum_{t=0}^n \frac{(C_t - N_t p_t)^2}{N_t p_t (1 - p_t)}, \quad p_t = q X_t,$$

is also explained.

In Section 5, the Pope, Nagai methods, and Ishioka and Kishida iteration method for VPA are explained. The simple VPA method using mortality rates instead of mortality coefficients is presented as follows:

$$N_{i+1} = (1 - D)N_i - C_i.$$

The transformation between mortality rates and mortality coefficients are

$$E = \frac{F}{F + M} (1 - e^{-F-M}), \quad D = \frac{M}{F + M} (1 - e^{-F-M}).$$

The Leslie matrix model related to VPA, and the linear programming using mortality rates are also presented.

Acknowledgments

The author thanks the late Mr. Fumihiko Kato and the late Mr. Kiyohide Ishioka for their encouragement. The author also thanks the two anonymous referees for their useful comments on earlier drafts of the manuscript.

References

- Aizawa Y, Takiguchi N. Consideration of the methods for estimating the age-composition from length frequency data with MS-Excel. *Bull. Jpn. Soc. Fish. Oceanogr.* 1999; **63**: 205–214 (in Japanese).
- Akamine T. A handy method for selecting and counting bivalve larvae at planktonic stage. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1981; **32**: 77–81 (in Japanese).
- Akamine T. A BASIC program to analyse the polymodal frequency distribution into normal distributions. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1982; **33**: 163–166 (in Japanese).
- Akamine T. The BASIC program to analyse the polymodal frequency distribution into normal distributions with Marquardt's method. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1984; **34**: 53–60 (in Japanese).
- Akamine T. Consideration of the BASIC programs to analyse the polymodal frequency distribution into normal distributions. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1985; **35**: 129–159 (in Japanese).
- Akamine T. Expansion of growth curves using a periodic function and BASIC programs by Marquardt's method. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1986; **36**: 77–107.
- Akamine T. A solution of the multi-cohort model by Marquardt's method. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1987a; **37**: 225–257.
- Akamine T. Comparison of algorithms of several methods for estimating parameters of a mixture of normal distributions. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1987b; **37**: 259–277.
- Akamine T. Evaluation of error caused by histogram on estimation of parameters for a mixture of normal distributions. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1988a; **38**: 171–185.
- Akamine T. Estimation of parameters for Richards model. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1988b; **38**: 187–200.
- Akamine T. An interval estimation for extraction using Bayesian statistics. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1989a; **39**: 9–17.
- Akamine T. An interval estimation for the Petersen method using Bayesian statistics. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1989b; **39**: 19–35.

- Akamine T. A new standard formula for seasonal growth of fish in population dynamics. *Nippon Suisan Gakkaishi* 1993; **59**: 1857–1863.
- Akamine T. Definition of the exponential function and its application for population dynamics. *Bull. Jpn. Soc. Fish. Oceanogr.* 1994; **58**: 317–320 (in Japanese).
- Akamine T. Mathematical approach to virtual population analysis. *Fish. Sci.* 1995a; **61**: 167.
- Akamine T. Relationship between Leslie matrix and cohort analysis. *Fish. Sci.* 1995b; **61**: 722.
- Akamine T. A mathematical study of fish growth formula in population dynamics. *Bull. Natl. Res. Inst. Fish. Sci.* 1995c; **7**: 189–263 (in Japanese).
- Akamine T. Introduction to cohort analysis (VPA). *Bull. Jpn. Soc. Fish. Oceanogr.* 1995d; **59**: 424–437 (in Japanese).
- Akamine T. Discrete fishing equations and linear programming. *Bull. Jpn. Soc. Fish. Oceanogr.* 1996; **60**: 252–258 (in Japanese).
- Akamine T. Optimum control theory for the dynamic pool model. *Bull. Natl. Res. Inst. Fish. Sci.* 1997; **10**: 135–167 (in Japanese).
- Akamine T. Convergence of Hasselblad method for estimating proportions of a mixture of probability distributions. *Fish. Sci.* 1998; **64**: 997–998.
- Akamine T. Convergence of Hasselblad method for estimating parameters of a mixture of normal distributions. *Bull. Natl. Res. Inst. Fish. Sci.* 1999; **14**: 49–58 (in Japanese).
- Akamine T. Mathematical study for approximations and iterations of the virtual population analysis. *Bull. Natl. Res. Inst. Fish. Sci.* 2001; **16**: 1–16 (in Japanese).
- Akamine T. Comparison between conventional and Bayesian statistics in interval estimation for quadrat method and Petersen method. *Bull. Fish. Res. Agen.* 2002; **2**: 25–34 (in Japanese).
- Akamine T. Statistical test and model selection of fish growth formula. *Bull. Jpn. Soc. Fish. Oceanogr.* 2004; **68**: 44–51 (in Japanese).
- Akamine T. Mixture of normal distributions and EM algorithm. *Bull. Jpn. Soc. Fish. Oceanogr.* 2005; **69**: 174–183 (in Japanese).
- Akamine T. *Suisan Sigen Kaiseki no Kiso*. Kouseisha Kouseikaku, Tokyo. 2007; 115 pp.
- Akamine T, Matsumiya Y. Mathematical analysis of age–length key method for estimating age composition from length composition. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1992; **42**: 17–24.
- Akamine T, Kishino H, Hiramatsu K. Non-biased interval estimation of Leslie’s removal method. *Bull. Japan Sea Natl. Fish. Res. Inst.* 1992; **42**: 25–39.
- Aksland M. A general cohort analysis method. *Biometrics* 1994; **50**: 917–932.
- Appeldoorn RS. Modification of a seasonally oscillating growth function for use with mark–recapture data. *J. Cos. Int. Explor. Mer* 1987; **43**: 194–198.
- Awaya T. Two-dimensional curve fitting in counting experiments. *Nucl. Instr. Meth.* 1983; **212**: 311–317.
- Awaya T. *Data Kaiseki*. Gakkai Shuppan Center, Tokyo. 1991; 270 pp.
- Cloern JE, Nichols FH. A von Bertalanffy growth model with a seasonally varying coefficient. *J. Fish. Res. Board Can.* 1978; **35**: 1479–1482.
- Dark TA. Age and growth of Pacific hake, *Merluccius productus*. *Fish. Bull.* 1975; **73**: 336–355.
- France J, Thornley JHM. *Mathematical Models in Agriculture*. Butterworth, London. 1984; pp. 75–94.
- Gorie S. Estimation of the parameters in von Bertalanffy’s growth formula by MS-Excel. *Suisanzoshoku* 2001; **49**: 519–527 (in Japanese).
- Gorie S. Estimation of parameters in a mixture of normal distributions from length frequency composition and growth formula by MS-Excel. *Suisanzoshoku* 2002; **50**: 243–249 (in Japanese).
- Group of Statistics, Pelagic Resource Section, SEIKAI Regional Fisheries Research Laboratory. Data of the investigation of September to November in 1950. Report of the Investigations of Trawl Fishery in the China Sea, No. 2. SEIKAI Regional Fisheries Research Laboratory, Nagasaki, Japan. 1951; pp. 1–28.
- Haddon M. *Modelling and Quantitative Methods in Fisheries*. Chapman and Hall, New York. 2001; 406 pp.
- Hasselblad V. Estimation of parameters for a mixture of normal distributions. *Technometrics* 1966; **8**: 431–444.
- Hiramatsu K. A theoretical study of equations used in virtual population analysis. *Fish. Sci.* 1995; **61**: 752–754.
- Iri M. *Suuchi Keisan*. Asakura Publishing, Tokyo. 1981; 173 pp.
- Ishioka K, Kishida T. Studies on the algorithm for solving the catch equation and its accuracy. *Bull. Nansei Reg. Fish. Res. Lab.* 1985; **19**: 111–120 (in Japanese).

- Kamijoh Y, Yokomatsu Y, Andoh K. Hamaguri no sigen baiyou gijutu kaihatu kennkyuu houkokusho. Oita Prefectural Agriculture, Forestry and Fisheries Research Center. 1985; 54 pp.
- Kimura DK. Likelihood methods for the von Bertalanffy growth curve. *Fish. Bull.* 1980; **77**: 765–776.
- Kiso K, Akamine T, Ohnishi S, Matsumiya Y. Mathematical examinations of the growth of sea-run and fluvial forms of female Masu salmon *Oncorhynchus masou* in rivers of the southern Sanriku district, Honshu, Japan. *Nippon Suisan Gakkaishi* 1992; **58**: 1779–1784.
- Kuno E. *Doubutsu no Kotaigun Doutai Kenkyuu Hou 1, Kotaisuu Suitei Hou*. Kyoritsu Shuppan, Tokyo. 1986; 114 pp.
- Mangel M, Beder JH. Search and stock depletion: theory and applications. *Can. J. Fish. Aquat. Sci.* 1985; **42**: 150–163.
- Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.* 1963; **11**: 431–441.
- Megrey BA. Review and comparison of age-structured stock assessment models from theoretical and applied points of view. *American Fisheries Society Symposium* 1989; **6**: 8–48.
- Murata N, Ikeda S. Neural network to EM. In: Watanabe M, Yamaguchi K (eds). *EM Algorithm to Fukanzan Data no Shomondai*. Taga Shuppan, Tokyo. 2000; pp. 155–188.
- Nagai T. Cohort analysis nyuumon. *Enyou Suiken News* 1980; **38**: 10–12 (in Japanese).
- Nose Y. On the confidence limits corresponding to the estimate obtained by DeLury's logarithmic catch-effort method. *Nippon Suisan Gakkaishi* 1959; **24**: 953–956 (in Japanese).
- Ohnishi S, Akamine T. Extension of von Bertalanffy growth model incorporating growth patterns of soft and hard tissues in bivalve molluscs. *Fish. Sci.* 2006; **72**: 787–795.
- Pauly D, David N. ELEFAN 1, a BASIC program for the objective extraction of growth parameters from length-frequency data. *Meeresforsch.* 1981; **28**: 205–211.
- Pauly D, Gaschutz G. A simple method for fitting oscillating length growth data with a program for pocket calculators. I.C.E.S. CM 1979/G/24, Demersal Fish Committee. 1979; 26 pp.
- Pauly D, Soriano-Bartz M, Moreau J, Jarre-Teichmann A. A new model accounting for seasonal cessation of growth in fishes. *Aust. J. Mar. Freshwater Res.* 1992; **43**: 1151–1156.
- Pitcher TJ, MacDonald PDM. Two models for seasonal growth in fishes. *J. Appl. Ecol.* 1973; **10**: 599–606.
- Richards FJ. A flexible growth function for empirical use. *J. Exp. Bot.* 1959; **10**: 290–300.
- Sakamoto Y, Ishiguro M, Kitagawa G. *Akaike Information Criterion Statistics*. D.Reidel Publishing, Tokyo. 1986; xix+290 pp.
- Schnute J. A versatile growth model with statistically stable parameters. *Can. J. Fish. Aquat. Sci.* 1981; **38**: 1128–1140.
- Schnute J. A general theory for analysis of catch and effort data. *Can. J. Fish. Aquat. Sci.* 1985; **42**: 414–429.
- Somers IF. On a seasonally oscillating growth function. *Fishbite* 1988; **6**: 8–11.
- Takeuchi K, Fujino K. *2-Kou Bunpu to Poisson Bunpu*. University of Tokyo Press, Tokyo. 1981; 262 pp.
- Tanaka S. A method of analyzing the polymodal frequency distribution and its application to the length distribution of porgy. *Bull. Tokai Reg. Fish. Res. Lab.* 1956; **14**: 1–13 (in Japanese).
- Taylor CC. Growth equations with metabolic parameters. *J. Cons. Int. Explor. Mer* 1962; **27**: 270–286.
- Yamada S, Kitada S. *Seibutsu Sigen Toukeigaku*. Seizando, Tokyo. 1997; 263 pp.
- Zippin C. An evaluation of the removal method of estimating animal populations. *Biometrics* 1963; **12**: 163–169.